

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 4, 2005, 09:07:56 ; Search time 363.973 Seconds  
(without alignments)  
836.639 Million cell updates/sec

Title: US-09-017-715A-2\_COPY\_120\_127  
Perfect score: 41  
Sequence: 1 EEAQSGSD 8

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlh  
-O=/cg12.1/USPTO.spool.h/US090171715/runat.04052005.100744.25619/app\_query.fasta\_1.661  
-DB=EST -OPMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOPCPL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45  
-DOCALLGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US090171715\_@CGN\_1\_1\_5334\_@runat.04052005.100744.25619 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MAIN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :  
EST :  
1: gb\_esc1:\*  
2: gb\_esc2:\*  
3: gb\_hic:\*  
4: gb\_esc3:\*  
5: gb\_esc4:\*  
6: gb\_esc5:\*  
7: gb\_esc6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	41	100.0	233 2	BE171313 RCI-HT054
2	41	100.0	301 1	AA055968 zf22a02.8
3	41	100.0	327 1	AA056035 zf22a02.r
4	41	100.0	333 4	BM748818 K-EST0023
5	41	100.0	352 1	A1468480 t957912.x
6	41	100.0	368 5	BU729743 UI-E-CK1
7	41	100.0	378 1	AA293803 zt56h04.8
8	41	100.0	387 1	AA722407 z983h10.8
9	41	100.0	393 1	AA394097 zt56h04.1

C	10	41	100.0	404	7	CF529590	UI-1-BC1P
C	11	41	100.0	407	1	AV683707	AV683707
C	12	41	100.0	408	1	AA946606	0q38c12.8
C	13	41	100.0	412	1	A1139933	qa68b04.x
C	14	41	100.0	421	1	AV703171	AV703171
C	15	41	100.0	428	5	BX090816	BX090816
C	16	41	100.0	438	1	A1684600	wa64d12.x
C	17	41	100.0	438	1	A1936527	wt29a07..x
C	18	41	100.0	442	1	AA804675	0f44b01.8
C	19	41	100.0	451	4	BG826435	602750062
C	20	41	100.0	472	4	BM704200	UI-E-CK1
C	21	41	100.0	473	6	CA421283	UI-H-FG0
C	22	41	100.0	474	4	BM706956	UI-E-CQ0
C	23	41	100.0	480	4	BM658990	UI-E-CK1
C	24	41	100.0	504	5	BU728272	UI-E-CQ0
C	25	41	100.0	509	2	BE299889	600944691
C	26	41	100.0	510	2	BE298825	60119383
C	27	41	100.0	519	1	A1016464	0c78h06.8
C	28	41	100.0	538	4	BM695726	UI-E-CQ1
C	29	41	100.0	544	4	BM655098	UI-E-CQ1
C	30	41	100.0	553	1	AA872836	0h7610.8
C	31	41	100.0	555	5	BX474500	DKFZP686D
C	32	41	100.0	555	7	CV028548	7090_Full1
C	33	41	100.0	558	4	B1548891	603189023
C	34	41	100.0	568	4	BG708703	602674249
C	35	41	100.0	578	5	BP212912	BP212912
C	36	41	100.0	582	5	BP197662	BP197662
C	37	41	100.0	583	5	BP200612	BP200612
C	38	41	100.0	583	5	BP346497	BP346497
C	39	41	100.0	584	1	AV708933	AV708933
C	40	41	100.0	584	5	BP201686	BP201686
C	41	41	100.0	588	4	B1755243	603023269
C	42	41	100.0	592	6	CA432299	UI-H-DH1
C	43	41	100.0	593	6	CA443166	UI-H-DH1
C	44	41	100.0	614	5	BU730570	UI-E-C11
C	45	41	100.0	617	1	AA633976	ac33e03.8

## ALIGNMENTS

RESULT 1  
BE171313  
LOCUS BE171313 233 bp mRNA linear EST 21-JUN-2000  
DEFINITION RCI-HT0545-100300-012-g05 HT0545 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE171313  
VERSION BE171313.1 GI:8634039  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R., Nagai,M.A., da Silva,W., Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=cl2=RC1-HT0545-100  
300-012-g05k3=2000-03-10&4=1)

Seq primer: puc18 forward  
High quality sequence start: 4  
High quality sequence stop: 233.  
Location/Qualifiers

1. 233

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_1ib="HT0545"

/note="Organ: head neck; Vector: puc18; Site 1: Sma1;  
Site 2: Sma1; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

## ORIGIN

## Alignment Scores:

Pred. No.:	205	Length:	233
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x BE171313 (1-233)

OY 1 GIUGLUALAGInserG1yYasp 8

Db 93 GAGGAGCCCCAGAGTGGGAGAGC 116

RESULT 2

AA055968/c 301 bp mRNA linear EST 17-SEP-1996

LOCUS zf22a02.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone

DEFINITION IMAGE:377642 3', mRNA sequence.

ACCESSION AA055968

VERSION AA055968.1 GI:1548325

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 301)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,

Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,

Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,

Trevaastis, E., Waterston, R., Williamson, A., Wohlmann, P. and

Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.jnl.gov) for further information.

Seg primer: -40M3 fwd. from Amerham.

Location/Qualifiers

1. 301

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:1285898"

/db\_xref="taxon:9606"

/clone="IMAGE:377642"

/sex="unknown"

/dev\_stage="19 weeks"

/lab host="DH10B (ampicillin resistant)"  
/clone\_1ib="Soares fetal heart NbHH19W"

/note="Organ: heart; Vector: pRT3D (pharmacia) with a  
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
GTGTCACATCTGAGTGGAGCGCCGACATCTTTTCTTTTCTTTT 3'],  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pRT3 vector  
(pharmacia). Library went through one round of  
normalization to a cot = 5. Library constructed by  
M. Fatima Bonaldo. This library was constructed from the  
same fetus as the fetal lung library, Soares fetal lung  
NbHH19W."

## ORIGIN

## Alignment Scores:

Pred. No.:	266	Length:	301
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AA055968 (1-301)

OY 1 GIUGLUALAGInserG1yYasp 8

Db 301 GAGGAGCCCCAGAGTGGGAGAGC 278

RESULT 3

AA056035

LOCUS zf22a02.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone

DEFINITION IMAGE:377642 5', mRNA sequence.

ACCESSION AA056035

VERSION AA056035.1 GI:1548374

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 327)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,

Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,

Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,

Trevaastis, E., Waterston, R., Williamson, A., Wohlmann, P. and

Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.jnl.gov) for further information.

Seg primer: -28M3 rev2 from Amerham

High quality sequence stop: 296.

Location/Qualifiers

1. 327

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:1285898"

/db\_xref="taxon:9606"

/clone="IMAGE:377642"

/sex="unknown"

/dev\_stage="19 weeks"

/lab host="DH10B (ampicillin resistant)"

/clone\_1ib="Soares fetal heart NbHH19W"

/note="Organ: heart; Vector: pRT3D (pharmacia) with a

modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5']  
 TGTTCACCATCTGAAGCGAGCGCCGCGCATCTTTTCTTTTCTTTT 3'],  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pRT3 vector  
 (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by  
 M. Fatima Bonaldo. This library was constructed from the  
 same fetus as the fetal lung library, Soares fetal lung  
 NBHL19W."

## ORIGIN

## Alignment Scores:

Pred. No.:	290	Length:	327
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-017-715a-2\_COPY\_120\_127 (1-8) x AA056035 (1-327)

## QY

1 GluGlulAgInSergIyGlyASP 8  
 |||||  
 29 GAGGAGGCCAGAGTGGGGAGAC 52

## Db

Bm748818

## RESULT 4

Bm748818

## LOCUS

Bm748818

## ACCESSION

Bm748818

## VERSION

Bm748818

## KEYWORDS

Bm748818

## SOURCE

Bm748818

## ORGANISM

Bm748818

## REFERENCE

Bm748818

## AUTHORS

Bm748818

## TITLE

Bm748818

## JOURNAL

Bm748818

## COMMENT

Bm748818

## FEATURES

Bm748818

## source

1. .333  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S9SNU601-1-C11"  
 /sex="M"  
 /tissue\_type="Ascites"  
 /cell\_type="Epithelial"  
 /cell\_line="SKU-601"  
 /lab\_host="TOP10"  
 /clone\_id="S9SNU601"  
 /note="Organ: Stomach; Vector: pME18-FL3; Site: 1: XhoI;  
 Site 2: XhoI. The poly (A) + RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tobacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including SfiI  
 site by treatment of T4 RNA ligase and the first strand  
 cDNA was synthesized with Superscript II using SfiI  
 oligo-dT primer. After first strand synthesis, RNA was

degraded by NaOH treatment and cDNA was amplified by PCR  
 reaction. The PCR products were digested with SfiI and  
 cloned into DraIII- digested pME18-FL3 vector. The  
 obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10" by electroporation method.  
 The cDNA libraries constructed by this method are  
 full-length enriched cDNA library."

## ORIGIN

## Alignment Scores:

Pred. No.:	295	Length:	333
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-017-715a-2\_COPY\_120\_127 (1-8) x Bm748818 (1-333)

## QY

1 GluGlulAgInSergIyGlyASP 8  
 |||||  
 237 GAGGAGGCCAGAGTGGGGAGAC 260

## Db

A1468480/c

## RESULT 5

A1468480/c

## LOCUS

A1468480

## ACCESSION

A1468480

## VERSION

A1468480

## KEYWORDS

A1468480

## SOURCE

A1468480

## ORGANISM

A1468480

## REFERENCE

A1468480

## AUTHORS

A1468480

## TITLE

A1468480

## JOURNAL

A1468480

## COMMENT

A1468480

## FEATURES

A1468480

## source

1. .352  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2112934"  
 /sex="male"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_id="NCI-CGAP\_P228"  
 /note="Organ: prostate; Vector: pRT3D-Pac (Pharmacia)  
 with a modified polylinker; Plasmid DNA from the  
 normalized library NCI CGAP P22 was prepared, and ss  
 circles were made in vitro. Following HAP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (cloneids  
 985608-986759, 1101192-1101959, and 1217928-1220615).  
 Subtraction by Bento Soares and M. Fatima Bonaldo. "

```

Alignment Scores:
Pred. No.:      312          Length:      352
Score:          41.00        Matches:      8
Percent Similarity: 100.00%   Conservative: 0
Percent Local Similarity: 100.00% Mismatches:    0
Query Match:     100.00%     Indels:         0
DB:              1          Gaps:            0

US-09-017-715A-2_COPY_120_127 (1-8) x A168480 (1-352)
QY      1  GIUcIuaIagInSerGIyGlyASP 8
|||||
Db       295 GAGCAGGCCCCACAGTCGGCGAGAC 272

RESULT 6
LOCUS   BU729743/c                               368 bp      mRNA      linear      EST 09-OCT-2002
DEFINITION   UI-E-Ck1-efj-m-04-0-UI_s1 UI-E-Ck1 Homo sapiens cDNA clone
              UI-E-Ck1-efj-m-04-0-UI 3', mRNA sequence.
ACCESSION   BU729743
VERSION     BU729743.1  GI:23652933
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 368)
AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
           discovery
JOURNAL    Genome Res. 6 (9), 791-806 (1996)
MEDLINE    97044477
PUBMED     8889548
COMMENT    Contact: Soares, MB
           Coordinated Laboratory for Computational Genomics
           University of Iowa
           375 Newton Road, 4156 MEERF, Iowa City, IA 52242, USA
           Tel.: 319 335 8250
           Fax: 319 335 9565
           Email: bento-soares@uiowa.edu
           Tissue Procurement: Dr. Gregg Hageman
           cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
           cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
           Clone Sequencing by: Dr. M. Bento Soares, University of Iowa
           DNA Distribution: Researchers may obtain clones from Research
           Genetics (www.resgen.com).
           Seq primer: ML3 FORWARD
           POLYA=yes.
FEATURES             Location/Qualifiers
     source            1..368
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-Ck1-efj-m-04-0-UI"
                     /tissue_type="Retina Foveal and Macular"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-Ck1"
                     /note="Organ: eye; Vector: pUT73-Pac (Pharmacia) with a
                           modified polylinker; Site 1: EcoR I; Site 2: Not I;
                           UI-E-Ck1 is a normalized cDNA library containing the
                           following tissue(s): Retina Foveal and Macular. The
                           library was constructed according to Bonaldo, Lennon and
                           Soares, Genome Research, 6:791-806, 1996. First strand
                           cDNA synthesis was primed with an oligo-dT primer
                           containing a Not I site. Double stranded cDNA was ligated
                           to an EcoR I adaptor, digested with Not I, and cloned
                           directionally into pUT73-Pac vector. The oligonucleotide
                           used to prime the synthesis of first-strand cDNA contains
                           a library tag sequence that is located between the Not I
                           site and the (dT)18 tail. The sequence tag for this
                           library is GTCC. This library was created for the program

```

```

Gene Discovery in the Visual System, supported by National
Eye Institute (NEI).
TAG_TISSUE=Foveal and Macular Retina
TAG_LIB=UI-E-CK1
TAG_SEQ=GTCC"

ORIGIN

Alignment Scores:

Pred. No.:      327          Length:      368
Score:          41.00        Matches:        8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:      100.00%      Indels:    0
DB:                5          Gaps:         0

US-09-017-715A-2_COPY_120_127 (1-8) x B0729743 (1-368)

Oy      1  GluGluAaGInSerGlyGlyAsp 8
      |||||
Db      320 GAGGAGGCCCCAGACTGGGGGAGAC 297

RESULT 7
AA293803          378 bp      mRNA      linear      EST 12-AUG-1997
LOCUS      z56h04.f1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
DEFINITION IMAGE:726391 3', mRNA sequence.
ACCESSION  AA293803
VERSION     AA293803.1  GI:1941726
KEYWORDS    EST.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 378)
AUTHORS    Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Joet,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
            Moore,B., Schellenberg,K., Seepce,W., Tan,F., Theising,B.,
            White,Y., Wyllie,T., Waterston,R. and Wilson,R.
            Washu-Merck EST Project 1997
            Unpublished (1997)
TITLE       Contact: Wilson RK
JOURNAL     Washington University School of Medicine
            444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LINT ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert length: 484      Std Error: 0.00
            Seq primer: -41m13 fwd, EF from Amersham.
            Location/Qualifiers
                1..378
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="GDB:5938504"
                /db_xref="taxon:9606"
                /clone="IMAGE:726391"
                /sex="Female"
                /tissue_type="ovarian tumor"
                /lab_host="DH10B (ampicillin resistant)"
                /clone_lib="Soares ovary tumor NbHOT"
                /note="Organ: ovary; Vector: p7773 (Pharmacia) with a
                modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                strand cDNA was primed with a Not I - oligo(dT) primer [5'
                TGTTACCAACTGAAAGTGGAGCGCGCGGCTTTTCTTTTCTTTT 3'],
                double-stranded cDNA was size selected, ligated to Eco RI
                adapters (Pharmacia), digested with Not I and cloned into
                the Not I and Eco RI sites of a modified p7773 vector
                (Pharmacia). Library constructed by Bento Soares and
                M.Fatima Bonaldo. "

ORIGIN

Alignment Scores:

```



Pred. No.: 336 Length: 378  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 1 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AA293803 (1-378)

QY 1 GluGluaIaGInSergIyGlyAsp 8  
 Db 302 GAGGAGGCCCACTGGGAGAGAC 279

RESULT 8  
 AA722407 387 bp mRNA linear EST 02-JAN-1998  
 LOCUS z983h10.g1 Soares fetal heart NBH19W Homo sapiens cDNA clone  
 DEFINITION IMAGE:400003 3', mRNA sequence.  
 ACCESSION AA722407  
 VERSION AA722407.1 GI:2740114  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 387)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Joet, S.,  
 Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,  
 Martin, J., Moore, B., Scheinberg, K., Stepien, M., Tan, F.,  
 Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.  
 WashU-NCI human EST Project  
 TITLE Unpublished (1997)  
 JOURNAL  
 COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 Seq primer: -40m13 fwd. ET from Amersham  
 High quality sequence stop: 384.

FEATURES  
 source  
 1..387  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:1307826"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:400003"  
 /sex="unknown"  
 /dev\_stage="19 weeks"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_1lb="Soares fetal heart NBH19W"  
 /note="Organ: heart; Vector: pT773D (Pharmacia) with a  
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTTCACCAATCTAGAGTGGAGCGCGGCACTTTTCTTTTCTTTT 3'],  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT773 vector  
 (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by  
 M.Fatima Bernaldo. This library was constructed from the  
 same fetus as the fetal lung library, Soares fetal lung  
 NBH19W."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 344 Length: 387  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0

DB: 1 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AA722407 (1-387)

QY 1 GluGluaIaGInSergIyGlyAsp 8  
 Db 297 GAGGAGGCCCACTGGGAGAGAC 274

RESULT 9  
 AA394097 393 bp mRNA linear EST 12-AUG-1997  
 LOCUS z156h04.r1 Soares ovary tumor NBH07 Homo sapiens cDNA clone  
 DEFINITION IMAGE:726391 5', similar to TR:G971580 G971580 SENSOR NEURON  
 SYNUCLEIN.1, mRNA sequence.  
 ACCESSION AA394097  
 VERSION AA394097.1 GI:2047067  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 393)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Joet, S.,  
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,  
 Moore, B., Scheinberg, K., Stepien, M., Tan, F., Theising, B.,  
 White, Y., Wyllie, T., Waterston, R. and Wilson, R.  
 WashU-Merck EST Project 1997  
 TITLE Unpublished (1997)  
 JOURNAL  
 COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 Insert length: 484 Std Error: 0.00  
 Seq primer: -28m13 rev2 ET from Amersham.

FEATURES  
 source  
 1..393  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:5938504"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:726391"  
 /sex="female"  
 /tissue\_type="ovarian tumor"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_1lb="Soares ovary tumor NBH07"  
 /note="Organ: ovary; Vector: pT773D (Pharmacia) with a  
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTTCACCAATCTAGAGTGGAGCGCGGCGGCTTTTCTTTTCTTTT 3'],  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT773 vector  
 (Pharmacia). Library constructed by Bento Soares and  
 M.Fatima Bernaldo."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 349 Length: 393  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AA394097 (1-393)

QY 1 GluGluaIaGInSergIyGlyAsp 8  
 Db 111 GAGGAGGCCCACTGGGAGAGAC 134

RESULT 10  
LOCUS CFS29590/c 404 bp mRNA linear EST 11-SEP-2003  
DEFINITION UI-1-BCIP-asy-b-08-0-UI.s1 NCI CGAP\_P13 Homo sapiens cDNA clone  
ACCESSION CFS29590  
VERSION CFS29590  
KEYWORDS EST. GI:34579285  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncigap>  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cga@bbsr.nlm.nih.gov](mailto:cga@bbsr.nlm.nih.gov)  
Tissue Procurement: Dr. Steven Brown  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
Seq primer: M13 FORWARD  
POLYA=yes.

FEATURES  
source  
1..404  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-1-BCIP-asy-b-08-0-UI"  
/tissue\_type="Placenta"  
/dev\_stage="8-9 weeks"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP P13"  
/note="Organ: Placenta; Vector: pT73-Pac (Pharmacia) with  
a modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
NCI CGAP\_P13 is a subcloned cDNA library constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into pT73-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tags for this library are GA, AGGA. For  
additional information, contact: Bento Soares,  
[bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
TAG\_TISSUE=Placenta human 8 week  
TAG\_LIB=UI-1-BCIP  
TAG\_SEQ=GA"

ORIGIN  
Alignment Scores:  
Pred. No.: 359 Length: 404  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 7 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x CFS29590 (1-404)

QY 1 GUGUUAAGInserGlyasp 8  
|||||  
DB 321 GAGGAGCCCAAGTGGGAGAC 298

RESULT 11  
AV683707

LOCUS AV683707 407 bp mRNA linear EST 16-JAN-2002  
DEFINITION AV683707 GKC Homo sapiens cDNA clone GKCEPB06 5', mRNA sequence.  
ACCESSION AV683707  
VERSION AV683707.1 GI:10285570  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.,  
1 (bases 1 to 407)  
Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,  
Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,  
Hu, G., Gu, J., Chen, Z., and Han, Z.  
TITLE Insight into hepatocellular carcinogenesis at transcriptome level  
by comparing gene expression profiles of hepatocellular carcinoma  
with those of corresponding noncancerous liver  
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)  
JOURNAL 11752456  
MEDLINE 21625106  
PUBMED 11752456  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex. 45)  
Fax: 86-21-50801922  
Email: [hanzg@chgc.sh.cn](mailto:hanzg@chgc.sh.cn)  
This clone is available at CHGC in Shanghai.

FEATURES  
source  
1..407  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GKCEPB06"  
/tissue\_type="hepatocellular carcinoma"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/clone\_lib="GKC"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
XhoI"

ORIGIN  
Alignment Scores:  
Pred. No.: 362 Length: 407  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AV683707 (1-407)

QY 1 GUGUUAAGInserGlyasp 8  
|||||  
DB 169 GAGGAGCCCAAGTGGGAGAC 192

RESULT 12  
AA946606/c 408 bp mRNA linear EST 23-JUL-1998  
LOCUS AA946606  
DEFINITION c938c12.s1 NCI CGAP Kids Homo sapiens cDNA clone IMAGE:158850 3'  
ACCESSION AA946606  
VERSION AA946606  
KEYWORDS EST. GI:3110001  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncigap>  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL ac:  
www-bio.llnl.gov/bdip/image/image.html  
Insert Length: 812 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 359.

## FEATURES

source

```
1..408
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1588630"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_id="NCI CGAP Kid5"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' ACTGGAAGATCGCGCGCATATTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo. "
```

## ORIGIN

## Alignment Scores:

Pred. No.:	363	Length:	408
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AA346606 (1-408)

Qy 1 GUGUUAAGInserGIyGlyAap 8

Db 309 GAGGAGGCCAGAGTGGGAGAC 286

## RESULT 13

A1139933/c

LOCUS 412 bp mRNA linear EST 05-OCT-1998  
DEFINITION g668B04.X1 Soares\_fetal\_heart NBHL19W Homo sapiens cDNA clone IMAGE:1691887 3' similar to TR:015104 015104 BC5G1 PROTEIN. ; mRNA sequence.

## ACCESSION

A1139933

VERSION A1139933.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

## REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

This clone is available royally-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 584 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham.

Location/Qualifiers

1..412

FEATURES

source

## ORIGIN

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1691887"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_id="Soares_fetal_heart NBHL19W"
/note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCACCATCGAAGCGAGCGCGCATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M. Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHL19W."
```

US-09-017-715A-2\_COPY\_120\_127 (1-8) x A1139933 (1-412)

## Qy

1 GUGUUAAGInserGIyGlyAap 8

Db 303 GAGGAGGCCAGAGTGGGAGAC 280

## RESULT 14

AV703171

LOCUS 421 bp mRNA linear EST 09-OCT-2000  
DEFINITION AV703171 ADB Homo sapiens cDNA clone ADBCF03 5', mRNA sequence.

## ACCESSION

AV703171

VERSION AV703171.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

## REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2000)

Contact: Zengqiang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919 (ex.45)

Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

Location/Qualifiers

1..421

FEATURES

source

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="ADBBCF03"

/tissue\_type="Adrenal gland"

/dev\_stage="Adult"

/lab\_host="SOLR"

/clone\_id="ADB"

/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:

## ORIGIN XhoI"

## Alignment Scores:

Pred. No.: 375 Length: 421  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 1 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AV703171 (1-421)

Qy 1 GIUGLUALAGInserGIyAsp 8  
 DB 374 GAGGAGGCCACAGTGGGGAGAC 397

## RESULT 15

BX090816

428 bp mRNA linear EST 23-JAN-2003

LOCUS BX090816 Soares ovary tumor NBHOT Homo sapiens cDNA clone

DEFINITION IMAGP998P081781 ; IMAGE:726391, mRNA sequence.

ACCESSION BX090816

VERSION BX090816.1 GI:27822248

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 428) Ebert, L., Heil, O., Hennig, S., Neubert, P., Patsch, E., Peters, M.,

AUTHORS Radelof, U., Schneider, D. and Korn, B.

TITLE Human Unigeneset - RZPD3

JOURNAL Unpublished (2003)

COMMENT

Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPD; IMAGP998P081781.

RZPDLIB; I.M.A.G.E. cDNA Clone Collection:

Human Unigeneset - RZPD3 (RZPDLIB No.972)

http://www.rzpd.de/ClonCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs.

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 101

Fax: +49 30 32639 111

www.rzpd.de

This clone is available royalty-free from RZPD;

contact RZPD (clone@rzpd.de) for further information. Seq primer:

M13r, Primer sequence: TTTCAACACGAGAAACAGCTATGAC.

Location/Qualifiers

1. 428

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGP998P081781 ; IMAGE:726391"

/sex="Female"

/tissue\_type="ovarian tumor"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares ovary tumor NBHOT"

/note="Organ: ovary; Vector: p773D (Pharmacia) with a

modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer (5'

TGTTACCAATCGAAGTGGAGCGCGGTTTTTTTTTTTTTTT 3'),

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified p773 vector

(Pharmacia). Library constructed by Bento Soares and

M.Fatima Bonaldo."

ORIGIN

Alignment Scores:

Pred. No.: 381 Length: 428

Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x BX090816 (1-428)

Qy 1 GIUGLUALAGInserGIyAsp 8  
 DB 111 GAGGAGGCCACAGTGGGGAGAC 134

Search completed: May 4, 2005, 13:46:36  
 Job time : 364.973 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 4, 2005, 09:07:56 ; Search time 636.953 Seconds  
(without alignments)  
836.639 Million cell updates/sec

Title: US-09-017-715A-2\_COPY\_94\_107  
Perfect score: 70  
Sequence: 1. VVRKEDLRSPAPQ 14

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+g2n.model -DEV=xlh  
-O=/cg92.1/USPTO.epool.h/US09017715/runat\_04052005\_100744\_25619/asp\_query.fasta\_1.661  
-DB=EST -OPMT=fastap -SUFFIX=rcs -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcr -NORM=exc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09017715\_@CGN\_1\_1\_5334\_@runat\_04052005\_100744\_25619 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:  
1: gb\_esc1.\*  
2: gb\_esc2.\*  
3: gb\_hlc.\*  
4: gb\_esc3.\*  
5: gb\_esc4.\*  
6: gb\_esc5.\*  
7: gb\_esc6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	70	100.0	161	2	BE774132 MRI-UM000
2	70	100.0	283	4	BM693306 UI-E-CK1
3	70	100.0	333	4	BM748818 K-EST0023
4	70	100.0	369	1	AV708398 AV708398
5	70	100.0	387	1	AA722407 Z963H10.8
6	70	100.0	393	1	AA394097 zt56h04.x
7	70	100.0	407	1	AA683707 AV683707
8	70	100.0	412	1	AI139933 qae8b04.x
9	70	100.0	421	1	AV703171 AV703171

10	70	100.0	428	5	BX090816
C 11	70	100.0	438	1	AI684600
C 12	70	100.0	442	1	AA804675
13	70	100.0	451	4	BG826435
14	70	100.0	462	5	BX474511
15	70	100.0	467	6	CB107161
16	70	100.0	472	4	BM704200 UI-E-CK1
17	70	100.0	474	4	BM706956
C 18	70	100.0	480	4	BM68990
C 19	70	100.0	489	5	BP201709
C 20	70	100.0	504	5	BU728272
C 21	70	100.0	509	2	BE299889
C 22	70	100.0	510	2	BE298825
C 23	70	100.0	519	1	AI016464
C 24	70	100.0	538	4	BM695726
C 25	70	100.0	544	4	BM665098
C 26	70	100.0	555	5	BX474500
27	70	100.0	555	7	CV028548
28	70	100.0	558	4	BI548891
29	70	100.0	568	4	BG708703
30	70	100.0	578	5	BP212912
31	70	100.0	582	5	BP197662
32	70	100.0	583	5	BP200612
33	70	100.0	584	1	AV708933
34	70	100.0	584	5	BP201686
35	70	100.0	588	4	BI755243
C 36	70	100.0	592	6	CA443299
C 37	70	100.0	593	6	CA443166
C 38	70	100.0	614	5	BU730570
C 39	70	100.0	617	1	AA633976
40	70	100.0	617	4	BG707764
41	70	100.0	623	5	BP381244
42	70	100.0	641	1	AL712443
43	70	100.0	653	4	BI757131
44	70	100.0	659	7	CN410061
45	70	100.0	706	4	BG328738

ALIGNMENTS

RESULT 1  
BE774132  
LOCUS  
DEFINITION MRI-UM0009-220500-015-g07\_1 UM0009 Homo sapiens cDNA, mRNA  
ACCESSION BE774132  
VERSION BE774132.1 GI:10227839  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
REFERENCE  
AUTHORS  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.P., Matsunuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jorgensen,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=62-MR1-UM0009-220>)  
500-015-g01\_let3=2000-05-22&cl=1

Seq primer: puc 18 forward  
High quality sequence stop: 161.  
Location/Qualifiers

1. 161  
/organism="Homo sapiens"

/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"

/clone\_lib="UM0009"

/note="Organ: uterus; Vector: puc18; Site 1: SmaI; Site 2: SmaI. A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

## Alignment Scores:

Pred. No.:	0.0174	Length:	161
Score:	70.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-017-715a-2\_COPY\_94\_107 (1-14) x BE74132 (1-161)

Qy 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14

Db 69 GTGGTCCGCAAGAGACTTGAGCCATCTGCCCCCAACAG 110

## RESULT 2

## BM693306

LOCUS BM693306 283 bp mRNA linear EST 28-FEB-2002  
DEFINITION UI-E-CK1-afm-1-14-0-UI r1 UI-E-CK1 Homo sapiens cDNA clone

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEHRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@iowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
CDNA library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
Seq primer: M13 Reverse.

## FEATURES

1. 283  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

/clone="UI-E-CK1-afm-1-14-0-UI"  
/tissue\_type="Retina Foveal and Macular"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-CK1"  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-CK1 is a normalized cDNA library containing the following tissue(s): Retina Foveal and Macular. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTCC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

## ORIGIN

Alignment Scores:	0.0316	Length:	283
Pred. No.:	70.00	Matches:	14
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	4		

US-09-017-715a-2\_COPY\_94\_107 (1-14) x BM693306 (1-283)

Qy 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14

Db 229 GTGGTCCGCAAGAGACTTGAGCCATCTGCCCCCAACAG 270

## RESULT 3

LOCUS BM748818 333 bp mRNA linear EST 04-MAR-2002  
DEFINITION K-EST0023804 S9SNU601 Homo sapiens cDNA clone S9SNU601-1-C11 5',

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Contact: Kim YS  
Unpublished (2002)  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 1 row: C column: 11  
High quality sequence stop: 333.  
Location/Qualifiers

## FEATURES

1. 333  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S9SNU601-1-C11"  
/sex="M"  
/tissue\_type="Ascites"  
/cell\_type="Epithelial"



Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AA722407 (1-387)

Qy 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
|||||  
375 GTGGTGGCAGAGAGACTTGAGCCATCTGCCCCCAACAG 334

RESULT 6  
AA394097 393 bp mRNA linear EST 12-AUG-1997  
LOCUS z55h04.r1 Soares ovary tumor NBH09 Homo sapiens cDNA clone  
DEFINITION IMAGE:726391.5, similar to TR:G971580 G971580 SENSORY NEURON  
SYNCDLEIN. ;, mRNA sequence.

ACCESSION AA394097  
VERSION AA394097  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 393)  
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,  
Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theisling, B.,  
White, Y., Wylie, T., Waterston, R. and Wilson, R.  
WashU-Merck EST Project 1997  
Unpublished (1997)

TITLE  
JOURNAL  
COMMENT Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

FEATURES  
source  
1.393  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:5938504"  
/db\_xref="taxon:9606"  
/clone="IMAGE:726391"  
/sex="Female"  
/tissue\_type="ovarian tumor"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares ovary tumor NBH09"  
/note="Organ: ovary; Vector: pRT73D (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo (dT) primer [5'  
TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTCTTTTCTTTT 3']  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pRT73 vector  
(Pharmacia). Library constructed by Bento Soares and  
M.Fatima Bonaldo."

## ORIGIN

## Alignment Scores:

Pred. No.: 0 0.447 Length: 393  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AA394097 (1-393)

Qy 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14

DB 33 GTGGTGGCAGAGAGACTTGAGCCATCTGCCCCCAACAG 74  
|||||

RESULT 7  
AV683707 407 bp mRNA linear EST 16-JAN-2002  
LOCUS AV683707 GKC Homo sapiens cDNA clone GKCPB06 5', mRNA sequence.  
DEFINITION AV683707  
ACCESSION AV683707  
VERSION AV683707.1 GI:10285570  
KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

## AUTHORS

## TITLE

JOURNAL  
MEDLINE  
PUBMED  
11752456  
Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex. 45)  
Fax: 86-21-50801922  
Email: hanzg@hgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
source  
1.407  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GKCPB06"  
/tissue\_type="hepatocellular carcinoma"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/clone\_lib="GKC"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
XhoI"

## ORIGIN

## Alignment Scores:

Pred. No.: 0.0464 Length: 407  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AV683707 (1-407)

Qy 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
|||||  
94 GTGGTGGCAGAGAGACTTGAGCCATCTGCCCCCAACAG 135

## RESULT 8

## LOCUS

## DEFINITION

AV683707 412 bp mRNA linear EST 05-OCT-1998  
qae6b04.x1 Soares fetal heart NBH119W Homo sapiens cDNA clone  
IMAGE:1691887 3', similar to TR:O15104 O15104 BCSG1 PROTEIN. ;, mRNA  
sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



REFERENCE 1 (bases 1 to 412)

NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: [cgap@remail.nih.gov](mailto:cgap@remail.nih.gov)

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Insert Length: 584 Std Error: 0.00

Seq primer: -40ml3 fwd. RT from Amersham.

Location/Qualifiers

1. 412

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1691887"

/sex="unknown"

/dev\_stage="19 weeks"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_id="Soares\_fetal\_heart\_NBH19W"

/note="Organ: heart; Vector: pT73D (Pharmacia) with a  
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TGTTACCACTGACAGTGGAGCGCGCCATCTTTTCTTTTCTTTT 3']  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pT73 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by  
M. Fatima Bonaldo. This library was constructed from the  
same fetus as the fetal lung library, Soares fetal lung  
NBH19W."

# ORIGIN

## Alignment Scores:

Pred. No.:	0.047	Length:	412
Score:	70.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
		Gaps:	0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x A1139933 (1-412)

QY 1 ValValatrglysguaapleuargproseralaproglngin 14

Db 381 GTGGTGGCAAGAGGAGCTTGAGGCCATCTGCCCCCAACAG 340

RESULT 9 AV703171 421 bp mRNA linear EST 09-OCT-2000

LOCUS AV703171 ADB Homo sapiens cDNA clone ADBCF03 5', mRNA sequence.

DEFINITION AV703171

ACCESSION AV703171

VERSION AV703171.1 GI:10720500

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 421)

CONTACT: Zeguang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919 (ex.45)

Fax: 86-21-50801922

Email: [hanzg@chgc.sh.cn](mailto:hanzg@chgc.sh.cn)

This clone is available at CHGC in Shanghai.

Location/Qualifiers

1. 421

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="ADBCF03"

/tissue\_type="Adrenal gland"

/dev\_stage="Adult"

/lab\_host="SOLR"

/clone\_id="ADB"

/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
XhoI"

# ORIGIN

## Alignment Scores:

Pred. No.:	0.0481	Length:	421
Score:	70.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
		Gaps:	0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AV703171 (1-421)

QY 1 ValValatrglysguaapleuargproseralaproglngin 14

Db 296 GTGGTGGCAAGAGGAGCTTGAGGCCATCTGCCCCCAACAG 337

RESULT 10

LOCUS EX090816

DEFINITION EX090816 Soares ovary tumor NBHOT Homo sapiens cDNA clone

ACCESSION IMAG998P081781 ; IMAGE:726391, mRNA sequence.

VERSION EX090816

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 428)

CONTACT: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPD; IMAG998P081781.

RZPDLIB; I.M.A.G.E. cDNA Clone Collection;

Human Unigeneset - RZPD3 (RZPDLIB No.972)

<http://www.rzpd.de/CloneCards/cgi-bin/showlib.pl.cgi/response?libno=972> Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 101

Fax: +49 30 32639 111

www.rzpd.de

This clone is available royalty-free from RZPD;

contact RZPD ([clone@rzpd.de](mailto:clone@rzpd.de)) for further information. Seq primer:

M3r, Primer sequence: TTTCAACAGGAAACAGCTATGAC.

## FEATURES

source

1. 428

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAG998P081781 ; IMAGE:726391"

/sex="Female"

/tissue\_type="ovarian tumor"

/lab\_host="DH10B (ampicillin resistant)"

/clone.lib="Soares ovary tumor NBH07"  
 /note="Organ: ovary; Vector: pRT73D (Pharmacia) with a  
 modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
 strand cDNA was primed with a Not I - c150 (dT) primer (5'  
 TGTTCACATCTGAGTGGAGCGCGCGGTTTCTTTTCTTTT 3'),  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and Eco RI  
 the Not I and Eco RI sites of a modified pRT73 vector  
 (Pharmacia). Library constructed by Bento Soares and  
 M.Facima Bonaldo."

## ORIGIN

## Alignment Scores:

Pred. No.:	0.0489	Length:	428
Score:	70.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	5	Gaps:	0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x BK090816 (1-428)

Qy 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
 |||||  
 Db 33 GTGGTGGCGAAGAGAGACTTGAGCCATCTGCCCCCAACAG 74

## RESULT 11

AI684600 438 bp mRNA linear EST 07-MAR-2000  
 LOCUS wa84d12.x1 Soares\_NFL\_T\_GBC\_S1 Homo sapiens cDNA clone  
 DEFINITION IMAGE:2302871 3' similar to TR:015104 O15104 BCSG1 PROTEIN. [1] ;  
 mRNA sequence.

ACCESSION AI684600.1 GI:4895894  
 VERSION EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 438)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-remail.nih.gov

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

Insert Length: 505 Std Error: 0.00

Seg primer: -40UP from Gibco.

Location/Qualifiers

1.438

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2302871"

/lab\_host="DH10B"

/clone.lib="Soares\_NFL\_T\_GBC\_S1"

/note="Organ: pooled; Vector: pRT73D-Pac (Pharmacia) with

a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;

Equal amounts of plasmid DNA from three normalized

libraries (fetal lung NBH19W, testis NHT, and B-cell

NCI CGAP GCBI) were mixed, and ss circles were made in

vitro. Following HAP purification, this DNA was used as

tracer in a subtractive hybridization reaction. The driver

was PCR-amplified cDNAs from pools of 5,000 clones made

from the same 3 libraries. The pools consisted of

I.M.A.G.E. clones 297480-302087, 682632-687239,

726408-728711, and 729096-731399. Subtraction by Bento

Soares and M. Facima Bonaldo."

## ORIGIN

## Alignment Scores:

Pred. No.:	0.0502	Length:	438
Score:	70.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AI684600 (1-438)

Qy 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
 |||||  
 Db 378 GTGGTGGCGAAGAGAGACTTGAGCCATCTGCCCCCAACAG 337

## RESULT 12

LOCUS AA804675/c 442 bp mRNA linear EST 19-FEB-1998  
 DEFINITION oF44b01.s1 NCI\_CGAP CNS1 Homo sapiens cDNA clone IMAGE:1427017 3'  
 similar to SW:SYUN\_RAT Q63544 SENSORY NEURON SYNUCLEIN. ; mRNA  
 sequence.

ACCESSION AA804675.1 GI:2876076

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 442)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-remail.nih.gov

Insert Length: 795 Std Error: 0.00

Seg primer: -40m13 fwd. ET from Amersham

High quality sequence scop: 301.

Location/Qualifiers

1.442

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1427017"

/tissue\_type="substantia nigra"

/lab\_host="DH10B"

/clone.lib="NCI\_CGAP CNS1"

/note="Organ: Brain; Vector: pCMV-SPORT4; Site\_1: SalI;

Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.0 kb."

## ORIGIN

## Alignment Scores:

Pred. No.:	0.0506	Length:	442
Score:	70.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AA804675 (1-442)

Qy 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
 |||||  
 Db 391 GTGGTGGCGAAGAGAGACTTGAGCCATCTGCCCCCAACAG 350

## RESULT 13

LOCUS BG826435 451 bp mRNA linear EST 22-MAY-2001  
 DEFINITION 602750062F1\_NTH\_MGC\_17 Homo sapiens cDNA clone IMAGE:4903046 5',  
 mRNA sequence.

ACCESSION BG826435

VERSION BG826435.1 GI:14174022

KEYWORDS EST.

**SOURCE**  
ORGANISM Homo sapiens (human)

**REFERENCE**  
AUTHORS Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 451)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Plate: L1CM1800 row: K column: 15  
High quality sequence stop: 447.  
Location/Qualifiers

**FEATURES**  
source  
1..451  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4903046"  
/tissue\_type="rhabdomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1lb="NIH\_MGC\_17"  
/note="Organ: muscle; Vector: pOTB7; Site: 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

**ORIGIN**

Alignment Scores:  
Pred. No.: 0.0517 Length: 451  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715a-2\_COPY\_94\_107 (1-14) x BG826435 (1-451)

**Qy**  
1 ValValArlgLygsluApLeuAArgProSerAlaProGlnGln 14  
|||||  
Db 349 GTGGTGGCAAGAGGACCTTGAGGCCATGCCCCCAACAG 390  
|||||

**RESULT 14**  
BX474511 462 bp mRNA linear EST 04-SEP-2003  
LOCUS DKFZ686624170.1 686 (synonym: hlcc3) Homo sapiens cDNA clone  
DEFINITION DKFZ686624170.5', mRNA sequence.  
ACCESSION BX474511  
VERSION BX474511.1 GI:31668718  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 462)  
Bloeker, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., Osaenger, A., Pobo, G., Han, M. and Wiemann, S.  
EST (Bloeker, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., et al.)  
Unpublished (2003)  
Contact: MIPS  
MIPS  
Ingolstaedter Landstr.1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by GBR (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.  
No sl sequence available.  
This clone (DKFZ686624170) is available at the RZPD in Berlin. Please contact the RZPD: Resourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

**FEATURES**  
source  
1..462  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZ686624170"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1lb="686 (synonym: hlcc3)"  
/note="Vector: pTRIPLEX2; Site: 1: SfiIA; Site 2: SfiIB; cDNA collection"

**ORIGIN**

Alignment Scores:  
Pred. No.: 0.0531 Length: 462  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-09-017-715a-2\_COPY\_94\_107 (1-14) x BX474511 (1-462)

**Qy**  
1 ValValArlgLygsluApLeuAArgProSerAlaProGlnGln 14  
|||||  
Db 390 GTGGTGGCAAGAGGACCTTGAGGCCATGCCCCCAACAG 431  
|||||

**RESULT 15**  
CB107161 467 bp mRNA linear EST 28-JAN-2003  
LOCUS CB107161  
DEFINITION K-EST0145400 L3SN475 Homo sapiens cDNA clone L3SN475-14-A11 5', mRNA sequence.  
ACCESSION CB107161  
VERSION CB107161.1 GI:27932968  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 467)  
Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, D.M., Park, H.S., Kim, S. and Kim, Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 14 row: A column: 11  
High quality sequence stop: 467.  
Location/Qualifiers

**FEATURES**  
source  
1..467  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="L3SN475-14-A11"  
/sex="M"  
/cell\_line="SNU-475"  
/lab\_host="Top10F"  
/clone\_1lb="L3SN475"



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus model

Run on: May 4, 2005, 10:17:21 ; Search time 113.879 Seconds

(without alignments)  
748,404 Million cell updates/sec

Title: US-09-017-715A-2\_COPY\_94\_107

Perfect score: 70

Sequence: 1 VAKEDRPSAPQ 14

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5642217 seqs, 3043843248 residues

Total number of hits satisfying chosen parameters: 11284434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+g2n.model -DEV=xlh  
-Q=/cgn2\_1/USPRO.spool/h/US09017175/runat\_04052005\_100747\_25719/app.query.fasta\_1.661  
-DB=Published Applications NA -QPM=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=45 -DOCALL=OUTFMT=ptc -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MATELEN=2000000000 -USER=US09017175 @CGN\_1\_1.684 @runat\_04052005\_100747\_25719  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PTCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PTCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	100.0	384	19 US-10-826-157-5	Sequence 5, Appl
2	70	100.0	479	10 US-09-918-995-2705	Sequence 2705, Ap
3	70	100.0	550	9 US-09-954-531-613	Sequence 613, Ap
4	70	100.0	550	17 US-10-453-478-1	Sequence 1, Appl
5	70	100.0	550	17 US-10-843-641-1680	Sequence 1680, Ap
6	70	100.0	720	14 US-10-097-340-297	Sequence 297, Ap
7	70	100.0	720	17 US-10-282-174-469	Sequence 469, Ap
8	70	100.0	720	19 US-10-600-009-469	Sequence 469, Ap
9	70	100.0	796	9 US-09-925-298-171	Sequence 171, Ap
10	70	100.0	796	14 US-10-102-806-171	Sequence 171, Ap
11	53	75.7	478	9 US-09-925-298-172	Sequence 172, Ap
12	53	75.7	478	14 US-10-102-806-172	Sequence 172, Ap
13	53	75.7	4606	17 US-10-440-425-388	Sequence 388, Ap
14	53	75.7	5666	17 US-10-282-174-72	Sequence 72, Appl
15	53	75.7	5666	17 US-10-282-174-73	Sequence 73, Appl
16	53	75.7	5666	19 US-10-600-009-72	Sequence 72, Appl
17	53	75.7	5666	19 US-10-600-009-73	Sequence 73, Appl
18	53	75.7	6012	17 US-10-282-174-483	Sequence 483, App
19	53	75.7	6012	19 US-10-600-009-483	Sequence 483, App
20	48	68.6	135638	16 US-10-314-657-1	Sequence 1, Appl
21	48	68.6	135638	19 US-10-473-193-1	Sequence 1, Appl
22	47	67.1	657	17 US-10-282-122A-23679	Sequence 23679, A
23	46	65.7	794	18 US-10-425-115-165876	Sequence 165876,
24	45	64.3	435	18 US-10-425-115-141921	Sequence 141921,
25	45	64.3	465	9 US-09-864-761-6012	Sequence 6012, Ap
26	45	64.3	1077	13 US-10-027-632-101236	Sequence 101236,
27	45	64.3	1077	13 US-10-027-632-101237	Sequence 101237,
28	45	64.3	1077	17 US-10-027-632-101236	Sequence 101236,
29	45	64.3	1077	17 US-10-027-632-101237	Sequence 101237,
30	45	64.3	1435	17 US-10-282-122A-26255	Sequence 26255, A
31	45	64.3	1435	17 US-10-282-122A-26255	Sequence 26255, A
32	45	64.3	1473	17 US-10-282-122A-28635	Sequence 28635, A
33	45	64.3	1488	9 US-09-738-626-1425	Sequence 1425, Ap
34	45	64.3	1491	17 US-10-282-122A-17537	Sequence 17537, A
35	45	64.3	16010	17 US-10-085-117-262	Sequence 262, Appl
36	45	64.3	3309400	9 US-09-738-626-1	Sequence 1, Appl
37	44	62.9	395	18 US-10-425-115-118311	Sequence 118311,
38	44	62.9	661	15 US-10-259-165-660	Sequence 660, App
39	44	62.9	775	17 US-10-424-599-9676	Sequence 9676, Ap
40	44	61.4	412	18 US-10-425-115-52591	Sequence 52591, A
41	43	61.4	465	9 US-09-864-761-4889	Sequence 4889, Ap
42	43	61.4	547	16 US-10-029-386-9553	Sequence 9553, Ap
43	43	61.4	571	9 US-09-864-761-9722	Sequence 9722, Ap
44	43	61.4	800	18 US-10-363-345A-37659	Sequence 37659, A
45	43	61.4	800	18 US-10-363-345A-37660	Sequence 37660, A

## ALIGNMENTS

RESULT 1  
US-10-826-157-5  
Sequence 5, Application US/10826157  
Publication No. US20050064548A1  
GENERAL INFORMATION:  
APPLICANT: Lindquist, Susan L.  
TITLE OF INVENTION: YEAST ECTOPICALLY EXPRESSING ABNORMALLY  
FILE REFERENCE: 17481-003001  
CURRENT APPLICATION NUMBER: US/10/826,157  
PRIOR FILING DATE: 2004-04-16  
PRIOR APPLICATION NUMBER: US 60/472,317  
PRIOR FILING DATE: 2003-05-20  
PRIOR APPLICATION NUMBER: US 60/463,284  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 384

TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-826-157-5

Alignment Scores:  
Pred. No.: 0.00156 Length: 384  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 19 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-10-826-157-5 (1-384)

QY 1 ValValArGlyGluAspLeuArgProSerAlaProGln 14  
DB 280 GTGGTGGCAGAGAGACTTGAGGCCATCTGCCCAACAG 321

RESULT 2  
US-09-918-995-2705  
Sequence 2705, Application US/09918995  
Publication No. US20030073623A1  
GENERAL INFORMATION:  
APPLICANT: HySeq, Inc.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
FILE REFERENCE: 20411-756  
CURRENT APPLICATION NUMBER: US/09/918, 995  
CURRENT FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: US/09/235, 076  
PRIOR FILING DATE: 1999-01-20  
NUMBER OF SEQ ID NOS: 38054  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2705  
LENGTH: 479  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(479)  
OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-2705

Alignment Scores:  
Pred. No.: 0.00193 Length: 479  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-918-995-2705 (1-479)

QY 1 ValValArGlyGluAspLeuArgProSerAlaProGln 14  
DB 293 GTGGTGGCAGAGAGACTTGAGGCCATCTGCCCAACAG 334

RESULT 3  
US-09-954-531-613  
Sequence 613, Application US/09954531  
Patent No. US20020165180A1  
GENERAL INFORMATION:  
APPLICANT: Weaver, Zoe  
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cano  
FILE REFERENCE: 689290-77  
CURRENT APPLICATION NUMBER: US/09/954, 531  
CURRENT FILING DATE: 2002-05-02  
PRIOR APPLICATION NUMBER: US/60/233, 133  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US/60/234, 009  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234, 034

PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234, 509  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: US/60/234, 567  
PRIOR FILING DATE: 2000-09-22  
NUMBER OF SEQ ID NOS: 1392  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 613  
LENGTH: 550  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-954-531-613

Alignment Scores:  
Pred. No.: 0.00222 Length: 550  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-954-531-613 (1-550)

QY 1 ValValArGlyGluAspLeuArgProSerAlaProGln 14  
DB 291 GTGGTGGCAGAGAGACTTGAGGCCATCTGCCCAACAG 332

RESULT 4  
US-10-453-478-1  
Sequence 1, Application US/10453478  
Publication No. US20030208043A1  
GENERAL INFORMATION:  
APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,  
TITLE OF INVENTION: Human Genes, Sequences and  
Expression Products  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
CITY: ROSELAND  
STREET: 6 BECKER FARM ROAD  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/453, 478  
FILING DATE: 04-Jun-2003  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/705, 771  
FILING DATE: August 30, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-346 (PF196)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 973-994-1700  
TELEFAX: 973-994-1744  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 550 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-453-478-1

```
Alignment Scores:
Pred. No.: 0.00222 Length: 550
Score: 70.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-017-715a-2_COPY_94_107 (1-14) x US-10-453-478-1 (1-550)
Qy 1 ValValaArgLySGuaApLeuArgProSeRa1aProGIngin 14
Db 291 GTGCTGCCAAGAGAGACTTGAGGCCATCTGCCCCCAACAG 332

RESULT 5
US-10-843-641A-1680
; Sequence 1680, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1680
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-1680

Alignment Scores:
Pred. No.: 0.00222 Length: 550
Score: 70.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-09-017-715a-2_COPY_94_107 (1-14) x US-10-843-641A-1680 (1-550)
Qy 1 ValValaArgLySGuaApLeuArgProSeRa1aProGIngin 14
Db 291 GTGCTGCCAAGAGAGACTTGAGGCCATCTGCCCCCAACAG 332

RESULT 6
US-10-097-340-297
; Sequence 297, Application US/10097340
; Publication No. US20030087250A1

GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCHE
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 297
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-297

Alignment Scores:
Pred. No.: 0.00289 Length: 720
Score: 70.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-017-715a-2_COPY_94_107 (1-14) x US-10-097-340-297 (1-720)
Qy 1 ValValaArgLySGuaApLeuArgProSeRa1aProGIngin 14
Db 328 GTGCTGCCAAGAGAGACTTGAGGCCATCTGCCCCCAACAG 369

RESULT 7
US-10-282-174-469
; Sequence 469, Application US/10282174
; Publication No. US20030224380A1
; GENERAL INFORMATION:
; APPLICANT: Becker, Kenneth David
; APPLICANT: Velicelebi, Gonul
; APPLICANT: Elliot, Kathryn J.
; APPLICANT: Wang, Xin
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Bertam, Lars
; APPLICANT: Saunders, Aleister J.
; APPLICANT: Mullin, Kristina M.
; APPLICANT: Sampson, Andrew Johnson
; APPLICANT: Blacker, Deborah Lynne
```

TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10  
TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER  
FILE REFERENCE: 37481-3308  
CURRENT APPLICATION NUMBER: US 10/282,174  
PRIOR FILING DATE: 2002-10-25  
PRIOR APPLICATION NUMBER: US 60/339,525  
PRIOR FILING DATE: 2001-10-25  
PRIOR APPLICATION NUMBER: US 60/338,010  
PRIOR FILING DATE: 2001-11-08  
PRIOR APPLICATION NUMBER: US 60/336,929  
PRIOR FILING DATE: 2001-11-08  
PRIOR APPLICATION NUMBER: US 60/338,363  
PRIOR FILING DATE: 2001-11-09  
PRIOR APPLICATION NUMBER: US 60/337,052  
PRIOR FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: US 60/368,919  
PRIOR FILING DATE: 2002-03-28  
NUMBER OF SEQ ID NOS: 564  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 469  
LENGTH: 720  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: allele  
LOCATION: 30,57,85,243,250,377,512,531,555,561,672  
OTHER INFORMATION: N is any  
US-10-282-174-469  
Alignment Scores:  
Pred. No.: 0.00289 Length: 720  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 17  
US-09-017-715a-2\_copy\_94\_107 (1-14) x US-10-282-174-469 (1-720)  
Qy 1 ValValAArgLyGluAspLeuArgProSerAlaProGlnGln 14  
Db 328 GTGGTGGCGCAGAGAGACTTGAGGCCATCTGCCCCCAACAG 369  
RESULT 8  
US-10-600-009-469  
Sequence 469, Application US/10600009  
Publication No. US2005009031A1  
GENERAL INFORMATION:  
APPLICANT: Becker, Kenneth David  
APPLICANT: Velicelbi, Gonul  
APPLICANT: Elliot, Kathryn J.  
APPLICANT: Wang, Xin  
APPLICANT: Tanzi, Rudolph E.  
APPLICANT: Bertram, Lars  
APPLICANT: Saunders, Aleister J.  
APPLICANT: Mullin, Kristina M.  
APPLICANT: Sampson, Andrew Johnson  
APPLICANT: Blacker, Deborah Lynne  
TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10  
TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER  
FILE REFERENCE: 37481-3308  
CURRENT APPLICATION NUMBER: US/10/600,009  
CURRENT FILING DATE: 2003-06-18  
PRIOR APPLICATION NUMBER: US 60/339,525  
PRIOR FILING DATE: 2001-10-25  
PRIOR APPLICATION NUMBER: US 60/338,010  
PRIOR FILING DATE: 2001-11-08  
PRIOR APPLICATION NUMBER: US 60/336,929  
PRIOR FILING DATE: 2001-11-08  
PRIOR APPLICATION NUMBER: US 60/338,363  
PRIOR FILING DATE: 2001-11-09

PRIOR APPLICATION NUMBER: US 60/337,052  
PRIOR FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: US 60/368,919  
PRIOR FILING DATE: 2002-03-28  
PRIOR APPLICATION NUMBER: US 10/282,174  
PRIOR FILING DATE: 2002-10-25  
NUMBER OF SEQ ID NOS: 564  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 469  
LENGTH: 720  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: allele  
LOCATION: 30,57,85,243,250,377,512,531,555,561,672  
OTHER INFORMATION: N is any  
US-10-600-009-469  
Alignment Scores:  
Pred. No.: 0.00289 Length: 720  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 19  
US-09-017-715a-2\_copy\_94\_107 (1-14) x US-10-600-009-469 (1-720)  
Qy 1 ValValAArgLyGluAspLeuArgProSerAlaProGlnGln 14  
Db 328 GTGGTGGCGCAGAGAGACTTGAGGCCATCTGCCCCCAACAG 369  
RESULT 9  
US-09-925-298-171  
Sequence 171, Application US/09925298  
Publication No. US2002003976A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA103  
CURRENT APPLICATION NUMBER: US/09/925,298  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05881  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1998-03-12  
NUMBER OF SEQ ID NOS: 846  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 171  
LENGTH: 796  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-298-171  
Alignment Scores:  
Pred. No.: 0.00319 Length: 796  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 9  
US-09-017-715a-2\_copy\_94\_107 (1-14) x US-09-925-298-171 (1-796)  
Qy 1 ValValAArgLyGluAspLeuArgProSerAlaProGlnGln 14  
Db 388 GTGGTGGCGCAGAGAGACTTGAGGCCATCTGCCCCCAACAG 429  
RESULT 10  
US-10-102-806-171  
Sequence 171, Application US/10102806  
Publication No. US2003005442A1  
GENERAL INFORMATION:



```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 171
; LENGTH: 796
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-806-171

Alignment Scores:
Pred. No.: 0.00319      Length: 796
Score: 70.00           Matches: 14
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 14        Indels: 0
DB: 14                 Gaps: 0

US-09-017-715A-2_COPY_94_107 (1-14) x US-10-102-806-171 (1-796)

Cy 1 ValValAlrGlyGluApLeuArgProSerAlaProGlnGln 14
Db 388 GTGCTGCCGACAGAGGACTTGAAGCCATCTCCCCCAACAG 429

RESULT 11
US-09-925-298-172
; Sequence 172, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 172
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-298-172

Alignment Scores:
Pred. No.: 2.27      Length: 478
Score: 53.00         Matches: 10
Percent Similarity: 100.00%  Conservative: 1
Best Local Similarity: 90.91%  Mismatches: 0
Query Match: 75.71%  Indels: 0
DB: 9                Gaps: 0

US-09-017-715A-2_COPY_94_107 (1-14) x US-09-925-298-172 (1-478)

Cy 4 LyeGluApLeuArgProSerAlaProGlnGln 14
Db 80 CAGAGAGACTTGAAGCCATCTCCCCCAACAG 112

RESULT 12
US-10-102-806-172
; Sequence 172, Application US/10102806
; Publication No. US20030054421A1

GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 172
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-806-172

Alignment Scores:
Pred. No.: 2.27      Length: 478
Score: 53.00         Matches: 10
Percent Similarity: 100.00%  Conservative: 1
Best Local Similarity: 90.91%  Mismatches: 0
Query Match: 75.71%  Indels: 0
DB: 17               Gaps: 0

US-09-017-715A-2_COPY_94_107 (1-14) x US-10-102-806-172 (1-478)

Cy 4 LyeGluApLeuArgProSerAlaProGlnGln 14
Db 80 CAGAGAGACTTGAAGCCATCTCCCCCAACAG 112

RESULT 13
US-10-240-425-388
; Sequence 388, Application US/10240425
; Publication No. US2004003502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Schett, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 388
; LENGTH: 4606
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Genbank Accession No. US2004003502A1 AF044311
US-10-240-425-388

Alignment Scores:
Pred. No.: 21.1      Length: 4606
Score: 53.00         Matches: 10
Percent Similarity: 100.00%  Conservative: 1
Best Local Similarity: 90.91%  Mismatches: 0
Query Match: 75.71%  Indels: 0
DB: 17               Gaps: 0

US-09-017-715A-2_COPY_94_107 (1-14) x US-10-240-425-388 (1-4606)
```

```
Qy 4 LysGluAspLeuArgProSerAlaProGlnGln 14
Db 3950 CAGGAGACTTGAAGCCATCTGCCCCCAACAG 3982

RESULT 14
US-10-282-174-72
Sequence 72, Application US/10282174
Publication No. US20030224380A1
GENERAL INFORMATION:
APPLICANT: Becker, Kenneth David
APPLICANT: Vellicelebi, Goni
APPLICANT: Elliot, Kathryn J.
APPLICANT: Wang, Xin
APPLICANT: Tanzi, Rudolph E.
APPLICANT: Bertam, Lars
APPLICANT: Saunders, Aleister J.
APPLICANT: Mullin, Kristina M.
APPLICANT: Sampson, Andrew Johnson
APPLICANT: Blacker, Deborah Lynne
TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
FILE REFERENCE: 37481-3308
CURRENT APPLICATION NUMBER: US/10/282,174
CURRENT FILING DATE: 2002-10-25
PRIOR APPLICATION NUMBER: US 60/339,525
PRIOR FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: US 60/338,010
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: US 60/336,929
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: US 60/338,363
PRIOR FILING DATE: 2001-11-09
PRIOR APPLICATION NUMBER: US 60/337,052
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: US 60/368,919
PRIOR FILING DATE: 2002-03-28
NUMBER OF SEQ ID NOS: 564
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 72
LENGTH: 5666
TYPE: DNA
ORGANISM: Homo sapiens
US-10-282-174-72

Alignment Scores:
Pred. No.: 25.9 Length: 5666
Score: 53.00 Matches: 10
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 90.91% Mismatches: 0
Query Match: 75.71% Indels: 0
DB: 17 Gaps: 0

US-09-017-715A-2_COPY_94_107 (1-14) x US-10-282-174-72 (1-5666)
Qy 4 LysGluAspLeuArgProSerAlaProGlnGln 14
Db 4512 CAGGAGACTTGAAGCCATCTGCCCCCAACAG 4544

RESULT 15
US-10-282-174-73
Sequence 73, Application US/10282174
Publication No. US20030224380A1
GENERAL INFORMATION:
APPLICANT: Becker, Kenneth David
APPLICANT: Vellicelebi, Goni
APPLICANT: Elliot, Kathryn J.
APPLICANT: Wang, Xin
APPLICANT: Tanzi, Rudolph E.
APPLICANT: Bertam, Lars
APPLICANT: Saunders, Aleister J.
APPLICANT: Mullin, Kristina M.
```

```
APPLICANT: Sampson, Andrew Johnson
APPLICANT: Blacker, Deborah Lynne
TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
FILE REFERENCE: 37481-3308
CURRENT APPLICATION NUMBER: US/10/282,174
CURRENT FILING DATE: 2002-10-25
PRIOR APPLICATION NUMBER: US 60/339,525
PRIOR FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: US 60/338,010
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: US 60/336,929
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: US 60/338,363
PRIOR FILING DATE: 2001-11-09
PRIOR APPLICATION NUMBER: US 60/337,052
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: US 60/368,919
PRIOR FILING DATE: 2002-03-28
NUMBER OF SEQ ID NOS: 564
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 73
LENGTH: 5666
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 560,590,617,645,915,987,1723,1943,1950,3151,3178,3189,3284,
LOCATION: 4276,4311,4452,4995,5019,5025,5112,5136,5421,5648,5517
OTHER INFORMATION: N is any
FEATURE:
NAME/KEY: allele
LOCATION: 3779
OTHER INFORMATION: deletion: T
FEATURE:
NAME/KEY: allele
LOCATION: 4156
OTHER INFORMATION: insertion following nucleotide 4155
FEATURE:
NAME/KEY: allele
LOCATION: 4976
OTHER INFORMATION: deletion: C
US-10-282-174-73

Alignment Scores:
Pred. No.: 25.9 Length: 5666
Score: 53.00 Matches: 10
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 90.91% Mismatches: 0
Query Match: 75.71% Indels: 0
DB: 17 Gaps: 0

US-09-017-715A-2_COPY_94_107 (1-14) x US-10-282-174-73 (1-5666)
Qy 4 LysGluAspLeuArgProSerAlaProGlnGln 14
Db 4512 CAGGAGACTTGAAGCCATCTGCCCCCAACAG 4544

Search completed: May 4, 2005, 16:39:43
Job time : 126.546 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 4, 2005, 09:07:55 ; Search time 32.4161 Seconds

(without alignments)

706.682 Million cell updates/sec

Title: US-09-017-715A-2\_COPY\_94\_107

Perfect score: 70

Sequence: 1 VNRKEDRPSPAPQ 14

## Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Command line parameters:

-MODEL=frame+ p2n.model -DEV=xjh  
-Q=/cgn2\_1/USPRO.spool/h/US09017715/runat\_04052005\_100745\_25632/app\_query.fasta\_1.661  
-DB=Issued Patents NA -OPMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCTL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US09017715 -OCGN 1 116 -runat\_04052005\_100745\_25632 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

- Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Printed. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	100.0	550	3	US-08-705-771-1
2	70	100.0	550	4	US-08-417-540-1
3	70	100.0	702	4	US-09-949-016-1915
4	70	100.0	720	4	US-09-949-016-442
5	53	75.7	8607	4	US-09-949-016-13657
6	53	75.7	8608	4	US-09-949-016-12184
7	47	67.1	825	4	US-09-489-039A-172
8	45	64.3	4403765	3	US-09-103-840A-2
9	45	64.3	4411529	3	US-09-103-840A-1
10	43	61.4	1011	4	US-09-902-540-8804
11	43	61.4	1187	1	US-08-440-856A-2
12	43	61.4	10096	4	US-09-902-540-935

13	42	60.0	689	4	US-09-949-016-5443	Sequence 5443, Ap
14	42	60.0	690	3	US-09-419-568F-24	Sequence 24, Appl
15	42	60.0	690	3	US-09-354-243B-24	Sequence 24, Appl
16	42	60.0	1152	4	US-09-870-574-1	Sequence 1, Appl1
17	42	60.0	4797	3	US-09-419-568F-25	Sequence 25, Appl
18	42	60.0	4797	3	US-09-354-243B-25	Sequence 25, Appl
19	42	60.0	8888	4	US-09-949-016-17185	Sequence 17185, A
20	42	60.0	50950	4	US-09-949-016-16659	Sequence 16659, A
21	41.5	59.3	9005	4	US-09-902-540-7894	Sequence 7894, Ap
22	41.5	59.3	10793	4	US-09-902-540-1062	Sequence 1062, Ap
23	41	58.6	601	4	US-09-949-016-115274	Sequence 115274,
24	41	58.6	918	4	US-09-902-540-9560	Sequence 144544,
25	41	58.6	1073	1	US-08-356-405-8	Sequence 9560, Ap
26	41	58.6	1073	1	US-08-356-405-8	Sequence 8, Appl
27	41	58.6	1074	4	US-09-826-509-446	Sequence 446, App
28	41	58.6	1545	4	US-09-270-767-1178	Sequence 1178, App
29	41	58.6	1545	4	US-09-270-767-16460	Sequence 16460, A
30	41	58.6	2128	3	US-09-280-116-11	Sequence 10, Appl
31	41	58.6	2217	1	US-07-865-662F-9	Sequence 9, Appl
32	41	58.6	2217	3	US-08-374-219B-9	Sequence 9, Appl
33	41	58.6	2459	3	US-08-443-795-2	Sequence 2, Appl1
34	41	58.6	7386	4	US-09-949-016-13287	Sequence 13287, A
35	41	58.6	8268	1	US-08-375-709-10	Sequence 10, Appl
36	41	58.6	8268	1	US-08-752-929-10	Sequence 83, Appl
37	41	58.6	8268	4	US-09-231-899-83	Sequence 1086, Ap
38	41	58.6	15172	4	US-09-902-540-1086	Sequence 1161, A
39	41	58.6	15661	4	US-09-949-016-13161	Sequence 15660, A
40	41	58.6	18200	4	US-09-949-016-15660	Sequence 15661, A
41	41	58.6	18200	4	US-09-949-016-15661	Sequence 15662, A
42	41	58.6	19719	4	US-09-949-016-15662	Sequence 15663, A
43	41	58.6	19719	4	US-09-949-016-15663	Sequence 1, Appl
44	41	58.6	37895	1	US-08-375-709-1	Sequence 1, Appl
45	41	58.6	37895	1	US-08-752-929-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-08-705-771-1  
Sequence 1, Application US/08705771  
Patent No. 6054289  
GENERAL INFORMATION:  
APPLICANT: Paul Moore, Reiner Genz, Hongjin Ji,  
Jian Ni and Jing-Shan Hu  
TITLE OF INVENTION: Human Genes, Sequences and  
TITLE OF INVENTION: Expression Products  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
ADDRESS: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/705,771  
FILING DATE: August 30, 1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-346 (PFI96)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 973-994-1744  
TELEFAX: 973-994-1744  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 550 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-705-771-1

Alignment Scores:  
Pred. No.: 0.00135 Length: 550  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-08-705-771-1 (1-550)

Qy 1 ValValAArgLyGluAspLeuArgProSerAlaProGlnGln 14  
Db 291 GTGGTGGCAGAGAGAGACTTGAGGCCATCTGCCCCCAACAG 332

RESULT 2  
US-09-417-540-1  
Sequence 1, Application US/09417540  
Patent No. 6639052  
GENERAL INFORMATION:  
APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,  
Jian Ni and Jing-Shan Hu  
TITLE OF INVENTION: Human Genes, Sequences and  
Expression Products  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
CROCCI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/417,540  
FILING DATE: 14-Oct-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/705,771  
FILING DATE: August 30, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-346 (PFI96)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 973-994-1700  
TELEFAX: 973-994-1744  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 550 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-417-540-1

Alignment Scores:  
Pred. No.: 0.00135 Length: 550  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
DB: 0

Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-417-540-1 (1-550)

Qy 1 ValValAArgLyGluAspLeuArgProSerAlaProGlnGln 14  
Db 291 GTGGTGGCAGAGAGAGACTTGAGGCCATCTGCCCCCAACAG 332

RESULT 3  
US-09-949-016-1915  
Sequence 1915, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1915  
LENGTH: 702  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-1915

Alignment Scores:  
Pred. No.: 0.00178 Length: 702  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-949-016-1915 (1-702)

Qy 1 ValValAArgLyGluAspLeuArgProSerAlaProGlnGln 14  
Db 327 GTGGTGGCAGAGAGAGACTTGAGGCCATCTGCCCCCAACAG 368

RESULT 4  
US-09-949-016-442  
Sequence 442, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 442  
LENGTH: 720  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-442

## Alignment Scores:

Pred. No.: 0.00184 Length: 720  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-949-016-442 (1-720)

Qy 1 ValValArgLysGluAspLeuArgProSerAlaProGlnGln 14

Db 328 GTGGTCCGCAAGAGAGACTTGAGCCATCTGCCCCCAACAG 369

## RESULT 5

US-09-949-016-13657  
; Sequence 13657, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13657  
; LENGTH: 8607  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(8607)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13657

## Alignment Scores:

Pred. No.: 31.2 Length: 8607  
Score: 53.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 90.91% Mismatches: 0  
Query Match: 75.71% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-949-016-13657 (1-8607)

Qy 4 LysGluAspLeuArgProSerAlaProGlnGln 14

Db 5951 CAGAGAGACTTGAGCCATCTGCCCCCAACAG 5983

## RESULT 6

US-09-949-016-12184  
; Sequence 12184, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12184  
; LENGTH: 8608  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(8608)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12184

Qy 3 ArgLysGluAspLeuArgProSerAlaProGlnGln 14

Db 234 CGATTGAGATCTCCGCCCGCTGGCGCACGCCAG 199

## RESULT 8

US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: PRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-949-016-12184 (1-8608)

Qy 4 LysGluAspLeuArgProSerAlaProGlnGln 14

Db 5951 CAGAGAGACTTGAGCCATCTGCCCCCAACAG 5983

## RESULT 7

US-09-489-039A-172/c  
; Sequence 172, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489, 039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 172  
; LENGTH: 825  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-172

Qy 3 ArgLysGluAspLeuArgProSerAlaProGlnGln 14

Db 234 CGATTGAGATCTCCGCCCGCTGGCGCACGCCAG 199

## RESULT 8

US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: PRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-489-039A-172 (1-825)

Qy 3 ArgLysGluAspLeuArgProSerAlaProGlnGln 14

Db 234 CGATTGAGATCTCCGCCCGCTGGCGCACGCCAG 199

## RESULT 8

US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: PRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 4403765  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
FEATURE:  
OTHER INFORMATION: CDC 1551  
OTHER INFORMATION: "n" bases at various positions throughout the sequence  
OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Alignment Scores:  
Pred. No.: 7.98e+05 Length: 4403765  
Score: 45.00 Matches: 9  
Percent Similarity: 78.57% Conservative: 2  
Best Local Similarity: 64.29% Mismatches: 3  
Query Match: 64.29% Indels: 0  
DB: 3 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-103-840A-2 (1-4403765)

Oy 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
Db 3343973 GGTTCGCGTGGCGAGACCTGCTGCGCGCGACGCGCGCGACG 3343932

RESULT 9  
US-09-103-840A-1/C  
Sequence 1, Application US/09103840A  
Patent No. 6294328  
GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
FILE REFERENCE: 24366-20007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
CURRENT FILING DATE: 1998-06-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 4411529  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: H37RV  
US-09-103-840A-1

Alignment Scores:  
Pred. No.: 7.99e+05 Length: 4411529  
Score: 45.00 Matches: 9  
Percent Similarity: 78.57% Conservative: 2  
Best Local Similarity: 64.29% Mismatches: 3  
Query Match: 64.29% Indels: 0  
DB: 3 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-103-840A-1 (1-4411529)

Oy 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
Db 3349641 GGTTCGCGTGGCGAGACCTGCTGCGCGCGACGCGCGCGACG 3349600

RESULT 10  
US-09-902-540-8804  
Sequence 8804, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.

APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15889)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 8804  
LENGTH: 1011  
TYPE: DNA  
ORGANISM: Myxococcus xanthus  
US-09-902-540-8804

Alignment Scores:  
Pred. No.: 161 Length: 1011  
Score: 43.00 Matches: 7  
Percent Similarity: 84.62% Conservative: 4  
Best Local Similarity: 53.85% Mismatches: 2  
Query Match: 61.43% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-902-540-8804 (1-1011)

Oy 2 ValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
Db 46 ATGCCGAGCCCGAGCGTGGCGCGCGCGCGCGCGCGAGCGCG 84

RESULT 11  
US-08-440-856A-2  
Sequence 2, Application US/08440856A  
Patent No. 5750873  
GENERAL INFORMATION:  
APPLICANT: DELAPORTA, STEPHEN L.  
TITLE OF INVENTION: MATERIALS AND METHODS FOR PRODUCING  
NUMBER OF SEQUENCES: 9  
TITLE OF INVENTION: PLANTS WITH SINGLE-SEX FLOWERS  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 PENNSYLVANIA AVE. N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,856A  
FILING DATE: 15-MAY-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: MILLMAN, ROBERT A.  
REGISTRATION NUMBER: 36,217  
REFERENCE/DOCKET NUMBER: 05463-20001.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1517  
TELEFAX: (202) 887-0763  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1187 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-440-856A-2

Alignment Scores:  
Pred. No.: 193 Length: 1187  
Score: 43.00 Matches: 8  
Percent Similarity: 78.57% Conservative: 3

Best Local Similarity: 57.14% Mismatches: 3  
Query Match: 61.43% Indels: 0  
DB: 1 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-08-440-856A-2 (1-1187)

OY 1 ValValArgLysGluApLeuArgProSerAlaProGlnGln 14  
DB 467 GTCTTCGACCGCCGGGAGGTTGACCGCTCTCCGCTCA 508

RESULT 12  
US-09-902-540-935/C  
Sequence 935, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(115849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 935  
LENGTH: 10096  
TYPE: DNA  
ORGANISM: Myxococcus xanthus  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(10096)  
OTHER INFORMATION: unsure at all n locations  
US-09-902-540-935

Alignment Scores:  
Pred. No.: 2.19e+03 Length: 10096  
Score: 43.00 Matches: 7  
Percent Similarity: 84.62% Conservative: 4  
Best Local Similarity: 53.85% Mismatches: 2  
Query Match: 61.43% Indels: 0  
Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-902-540-935 (1-10096)

OY 2 ValArgLysGluApLeuArgProSerAlaProGlnGln 14  
DB 2820 ATGGCGAAGCCCGGACGTGCGCCCGCGGCGCCCGGAGCGC 2782

RESULT 13  
US-09-949-016-5443  
Sequence 5443, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASES, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: C1001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5443  
LENGTH: 689  
TYPE: DNA  
ORGANISM: Human

US-09-949-016-5443

Alignment Scores:  
Pred. No.: 156 Length: 689  
Score: 42.00 Matches: 8  
Percent Similarity: 90.00% Conservative: 1  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 60.00% Indels: 0  
Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-949-016-5443 (1-689)

OY 3 ArgLysGluApLeuArgProSerAlaPro 12  
DB 156 AGGAGAGGACGCTGCGCCCATCAGCTCC 185

RESULT 14  
US-09-419-568F-24  
Sequence 24, Application US/09419568F  
Patent No. 631613  
GENERAL INFORMATION:  
APPLICANT: Dumoutier, Laure  
APPLICANT: Louhed, Jamila  
APPLICANT: Renaud, Jean-Christophe  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fac  
TITLE OF INVENTION: (TTFs) The Proteins Encoded, and Uses Thereof  
FILE REFERENCE: LUD 5543.2  
CURRENT APPLICATION NUMBER: US/09/419,568F  
CURRENT FILING DATE: 1999-10-18  
PRIOR APPLICATION NUMBER: US09/354,243  
PRIOR FILING DATE: 1999-07-16  
PRIOR APPLICATION NUMBER: US09/178,973  
PRIOR FILING DATE: 1998-10-26  
NUMBER OF SEQ ID NOS: 29  
SEQ ID NO 24  
LENGTH: 690  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
US-09-419-568F-24

Alignment Scores:  
Pred. No.: 156 Length: 690  
Score: 42.00 Matches: 8  
Percent Similarity: 90.00% Conservative: 1  
Best Local Similarity: 80.00% Mismatches: 1  
Query Match: 60.00% Indels: 0  
Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-419-568F-24 (1-690)

OY 3 ArgLysGluApLeuArgProSerAlaPro 12  
DB 157 AGGAGAGGACGCTGCGCCCATCAGCTCC 186

RESULT 15  
US-09-354-243B-24  
Sequence 24, Application US/09354243B  
Patent No. 6359117  
GENERAL INFORMATION:  
APPLICANT: Dumoutier, Laure  
APPLICANT: Louhed, Jamila  
APPLICANT: Renaud, Jean-Christophe  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fac  
TITLE OF INVENTION: (TTFs) The Proteins Encoded, and Uses Thereof  
FILE REFERENCE: LUD 5543.1  
CURRENT APPLICATION NUMBER: US/09/354,243B  
CURRENT FILING DATE: 1999-07-16  
PRIOR APPLICATION NUMBER: US09/178,973  
PRIOR FILING DATE: 1998-10-26  
NUMBER OF SEQ ID NOS: 29  
SEQ ID NO 24

```

; LENGTH: 690
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-354-243B-24

```

```

Alignment Scores:
Pred. No.: 156 Length: 690
Score: 42.00 Matches: 8
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 60.00% Indels: 0
DB: 3 Gaps: 0

```

US-09-017-715A-2\_COPY\_94\_107 (1-14) x US-09-354-243B-24 (1-690)

```

QY 3 ArgLySGluAspLeuArgProSerAlaPro 12
   |||::||| ||||| ||||| |||||
DB 157 AGGAGAGAGAGAGCTGCGCCCATCAGCTCCC 186

```

Search completed: May 4, 2005, 09:42:48  
Job time : 667.416 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_p2n model

Run on: May 4, 2005, 09:07:52 ; Search time 102.322 Seconds  
(without alignments)  
809.955 Million cell updates/sec

Title: US-09-017-715A-2\_COPY\_94\_107

Perfect score: 70

Sequence: 1 VVRKEDRPAPQ 14

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959670667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+2pn.model -DEV=xjh  
-Q=/cgn2\_1/USPTO.spool.h/US09017715/runat\_04052005\_100743\_25600/app\_query.fasta\_1.661  
-DB=N\_Geneseq -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45  
-DOCALL=200 -THR\_SCORE=ppct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US09017715\_0CGN\_1\_1\_703\_0runat\_04052005\_100743\_25600 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N\_Geneseq\_16Dec04:\*

1: geneseqn1980s:\*\n2: geneseqn1990s:\*\n3: geneseqn2000s:\*\n4: geneseqn2001as:\*\n5: geneseqn2001bs:\*\n6: geneseqn2002as:\*\n7: geneseqn2002bs:\*\n8: geneseqn2003as:\*\n9: geneseqn2003bs:\*\n10: geneseqn2003cs:\*\n11: geneseqn2003ds:\*\n12: geneseqn2004as:\*\n13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	100.0	479	9	ACH15493 Human adu
2	70	100.0	488	12	ADM66887 Human hom
3	70	100.0	550	2	AAV42669 Human bre
4	70	100.0	550	3	AAA39470 Human HBG
5	70	100.0	550	6	ABL63343 Breast ca

6	70	100.0	550	6	ABV73813	ABV73813 Human gam
7	70	100.0	550	6	ABV73915	ABV73915 Human gam
8	70	100.0	550	10	ADG63568	Adg63568 Human amy
9	70	100.0	550	10	ADG47636	Adg47636 Human amy
10	70	100.0	720	2	AAX29997	Aax29997 Human per
11	70	100.0	720	6	AB876519	Ab876519 cDNA enco
12	70	100.0	720	10	AD843864	Ad843864 Human SNC
13	70	100.0	720	12	ADH54342	Adh54342 Human SNC
14	70	100.0	783	3	AAI93778	Aai93778 Human pol
15	70	100.0	796	3	AAF21784	Aaf21784 Human bre
16	70	100.0	990	13	ADR98806	Adr98806 Lung spec
17	54	77.1	677	2	AAK04876	Aak04876 Human gam
18	53	75.7	478	3	AAF21785	Aaf21785 Human bre
19	53	75.7	4606	6	ABT10161	Abt10161 Human bre
20	53	75.7	5666	10	AD843467	Ad843467 Human SNC
21	53	75.7	5666	10	AD843468	Ad843468 Polymorph
22	53	75.7	5666	12	ADH53945	Adh53945 Human SNC
23	53	75.7	5666	12	ADH53946	Adh53946 Human IDE
24	53	75.7	6012	10	AD843314	Ad843314 Human SNC
25	53	75.7	6012	12	ADH54356	Adh54356 Human SNC
26	49	70.0	787	2	AAT51183	Aat51183 Human bre
27	48	68.6	135638	10	ABX34289	Abx34289 S. atrool
28	47	67.1	657	8	ACA35809	Ac35809 Prokaryot
29	47	67.1	727	2	AAX29998	Aax29998 Mouse per
30	47	67.1	727	12	ADM66886	Adm66886 Murine ad
31	47	67.1	825	11	ACH94377	Ach94377 Klebsell
32	45	64.3	429	8	ACA01712	Ac01712 C. glutam
33	45	64.3	465	4	AAI16026	Aai16026 Probe #59
34	45	64.3	465	4	ABA58473	Ab58473 Human foe
35	45	64.3	465	4	AAI38127	Aai38127 Probe #68
36	45	64.3	465	4	ABA27546	Ab27546 Probe #60
37	45	64.3	465	4	AAK32274	Aak32274 Human bon
38	45	64.3	465	4	AAK06585	Aak06585 Human bra
39	45	64.3	465	4	ABS31980	Ab31980 Human liv
40	45	64.3	465	6	ABS07052	Ab07052 Human gen
41	45	64.3	1410	8	ACA37685	Ac37685 Prokaryot
42	45	64.3	1435	8	ACA38395	Ac38395 Prokaryot
43	45	64.3	1473	5	ACA40765	Ac40765 Prokaryot
44	45	64.3	1488	8	AAH66390	Aah66390 C glutami
45	45	64.3	1491	8	ACA96667	Ac96667 Prokaryot

## ALIGNMENTS

RESULT 1	
ACH15493	
ID	ACH15493 standard; cDNA; 479 BP.
XX	
AC	ACH15493;
XX	
DT	13-OCT-2003 (first entry)
XX	
DE	Human adult brain cDNA #2705.
XX	
KM	Human; ss: sequencing by hybridisation; SHH; expressed sequence tag; EST;
XX	genome mapping; biodiversity; genetic disorder.
OS	Homo sapiens.
XX	
PN	US2003073623-A1.
XX	
PD	17-APR-2003.
XX	
PF	30-JUL-2001; 2001US-00918995.
XX	
PR	30-JUL-2001; 2001US-00918995.
XX	
PA	(DRMA/) DRMANAC R. T.
PA	(LABA/) LABAT I.
PA	(STAC/) STACHE-CRAIN B.
PA	(DICK/) DICKSON M. C.
XX	(JONE/) JONES L. W.

PI Drmanac RT, Labat I, Stache-Crain B, Dickeon MC, Jones LW;  
XX WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
XX as hybridization probes, as oligomers for PCR, for chromosome and gene  
XX mapping, in the recombinant production of protein, or in generating  
XX antisenese DNA or RNA.  
PS Claim 1; SEQ ID NO 2705; 44pp; English.

CC The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocId=20030073623  
XX

SQ Sequence 479 BP; 120 A; 118 C; 178 G; 56 T; 0 U; 7 Other;

Alignment Scores:  
Pred. No.: 0.00451 Length: 479  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x ACHI5493 (1-479)

OY 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
Db |||||  
293 GTGGTCCGCAAGAGACTTGAGGCCATCGCCCCCAACAG 334

RESULT 2  
ADM66887  
ID ADM66887 standard; DNA; 488 BP.

AC ADM66887;  
XX  
XX  
DT 03-JUN-2004 (first entry)

XX Human homologue of murine adipocyte specific gamma synuclein DNA Seq 20.

XX human; adipocyte specific; gene; ds: adipose tissue; anti-obesity;  
XX high mobility group I-C protein; HMG1-C; obesity; leptin; ob; diabetes;  
XX antidiabetic; hypertensive; cardiovascular disease; anorectic;  
XX antidiabetic; hypotensive; gamma synuclein.

XX Homo sapiens.

XX WO2004011618-A2.

XX 05-FEB-2004.

XX 29-JUL-2003; 2003WO-US023684.

XX 29-JUL-2002; 2002US-0398785P.  
XX 12-JUN-2003; 2003US-0478206P.

XX (HMG1-C) HMG1-C INC.

PI Chada K, Chouinard R, Ashar H, Sayed AMD;  
XX WPI: 2004-143846/14.  
XX P-PSDB; ADM67167.  
XX  
XX Identifying adipocyte specific genes, useful for treating obesity or  
XX diabetes, and for identifying drug targets, by differential gene  
XX expression analysis between adipose tissue or stromal vascular tissue of  
XX mice of different genotypes.  
PS Claim 11; SEQ ID NO 20; 91pp; English.

XX This invention relates to a novel method for identifying genes that are  
XX over-expressed in adipose tissue and as such it provides targets for anti-  
XX obesity pharmaceutical compositions. Specifically, it refers to a high  
XX mobility group I-C protein (HMG1-C) that is associated with obesity and  
XX is epistatic to leptin, furthermore, it refers to the ob gene where an  
XX autosomal recessive trait is linked to obesity and diabetes. The present  
XX invention describes performing differential gene expression analysis  
XX between the white adipose tissue (WAT) or stromal vascular tissue (SVT)  
XX of any two different mice selected from a group consisting of wild-type,  
XX HMG1-C -/-, ob/ob, or HMG1-C -/- ob/ob genotype mice. Accordingly, using  
XX this method novel nucleotides and the encoded proteins thereof were  
XX identified that are adipocyte specific, and as such can be used for  
XX preventing adipogenesis, diagnosing and treating diabetes, obesity,  
XX hypertension and cardiovascular disease, as well as screening for  
XX compounds that can modulate or prevent adipogenesis and treat diabetes or  
XX obesity. These compositions exhibit anorectic, antidiabetic and  
XX hypotensive activities. This polynucleotide sequence is a human homologue  
XX of a murine adipocyte specific DNA sequence of the invention.

SQ Sequence 488 BP; 127 A; 119 C; 176 G; 66 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 0.0046 Length: 488  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x ADM66887 (1-488)

OY 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
Db |||||  
291 GTGGTCCGCAAGAGACTTGAGGCCATCGCCCCCAACAG 332

RESULT 3  
AAV42669  
ID AAV42669 standard; cDNA; 550 BP.

AC AAV42669;

XX 09-NOV-1998 (first entry)

XX Human breast cancer specific gene 1 (BCSG1) cDNA.

XX Breast cancer specific gene 1; BCSG1; human; metastasis; diagnosis;  
XX therapy; genetic marker; ds.

XX Homo sapiens.

XX WO9833915-A1.

XX 06-AUG-1998.

XX 03-FEB-1998; 98WO-US001804.

XX 03-FEB-1997; 97US-0037080P.

XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Ji H, Rosen CA,  
 XX WPI; 1998-446811/38.  
 DR P-PSDB; AAM63123.  
 XX  
 PT New isolated human breast cancer specific gene - used to develop products  
 PT for the diagnosis, clinical management and treatment of breast cancer and  
 PT metastases.  
 XX  
 PS Claim 4; Fig 1; 73pp; English.  
 XX  
 CC This cDNA clone corresponds to the transcript of the newly identified  
 CC human breast cancer specific gene 1 (BCSG1), and includes an open reading  
 CC frame for a 14.2 kda protein (see AAM63123). It was isolated from a  
 CC breast cancer cDNA library following an EST search for novel genes  
 CC differentially expressed in breast cancer versus healthy breast tissue.  
 CC The clone is deposited at ATCC 97175 and ATCC 97856. A gradient and stage  
 CC -specific BCSG1 expression has been demonstrated from virtually no  
 CC detectable expression in normal or benign breast to low level and partial  
 CC expression in low grade in situ breast carcinoma and high expression in  
 CC infiltrating malignant breast carcinomas. BCSG1 is useful as a breast  
 CC cancer progression marker. Recombinant vectors and host cells useful for  
 CC recombinant production of BCSG1 polypeptides (including epitope-bearing  
 CC polypeptides) are provided. BCSG1 polynucleotides, polypeptides and  
 CC antibodies can be used for the detection of breast cancer cells or breast  
 CC cancer metastasis, and to develop methods for the clinical management and  
 CC treatment of breast cancer  
 XX  
 SQ Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 0.00526 Length: 550  
 Score: 70.00 Matches: 14  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-017-715a-2\_COPY\_94\_107 (1-14) x AAV42669 (1-550)

QY 1 ValValArgLysGluAspLeuArgProSerAlaProGlnGln 14  
 Db 291 GTGGTGGCAAGAGGACCTTGAGGCATCTGCCCCCAACAG 332

RESULT 4

AAA39470  
 ID AAA39470 standard; DNA, 550 BP.

AC AAA39470;

DT 24-AUG-2000 (first entry)

DE Human HBGBA67A DNA.

XX Human; ADA2; cytostatic; gene therapy; treatment; cancer;  
 KW amyloid-like protein; ss.

OS Homo sapiens.

FT Key Location/Qualifiers  
 FT CDS 12..395  
 FT /\*tag= a  
 FT /product= "HBGBA67"

PN US6054289-A.

PD 25-APR-2000.

PF 30-AUG-1996; 96US-00705771.

PR 30-AUG-1995; 95US-0002993P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Moore PA;  
 XX WPI; 2000-338491/29.  
 DR P-PSDB; AAY87779.  
 XX  
 PT New polynucleotide encoding human AD2 is useful for treating cancer and  
 PT for isolating cDNAs and genes having similar biological activity.  
 XX  
 PS Disclosure; Col 27-28; 54pp; English.  
 XX  
 CC This invention describes a novel polynucleotide (I) encoding human ADA2.  
 CC The products of the invention have cytostatic activity and can be used  
 CC for gene therapy. (I) is useful for treating cancer; as primers and  
 CC probes for isolating full length cDNA and genes having similar biological  
 CC activity. This sequence encodes a polypeptide derived from the human  
 CC HBGBA67X clone which is an amyloid-like protein found in breast tissue  
 XX  
 SQ Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 0.00526 Length: 550  
 Score: 70.00 Matches: 14  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0

US-09-017-715a-2\_COPY\_94\_107 (1-14) x AAA39470 (1-550)

QY 1 ValValArgLysGluAspLeuArgProSerAlaProGlnGln 14  
 Db 291 GTGGTGGCAAGAGGACCTTGAGGCATCTGCCCCCAACAG 332

RESULT 5

ABL63343  
 ID ABL63343 standard; DNA, 550 BP.

AC ABL63343;

DT 15-MAY-2002 (first entry)

DE Breast cancer related gene sequence SEQ ID NO:1680.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;  
 KW gene; ds.

OS Homo sapiens.

PN WO200194629-A2.

PD 13-DEC-2001.

PF 30-MAY-2001; 2001WO-US010838.

PR 05-JUN-2000; 2000US-0209473P.  
 PR 05-JUN-2000; 2000US-0209531P.  
 PR 18-SEP-2000; 2000US-0233133P.  
 PR 18-SEP-2000; 2000US-0233617P.  
 PR 20-SEP-2000; 2000US-0234009P.  
 PR 20-SEP-2000; 2000US-0234034P.  
 PR 20-SEP-2000; 2000US-0234052P.  
 PR 22-SEP-2000; 2000US-0234509P.  
 PR 22-SEP-2000; 2000US-0234567P.  
 PR 25-SEP-2000; 2000US-0234923P.  
 PR 25-SEP-2000; 2000US-0234924P.  
 PR 25-SEP-2000; 2000US-0235077P.  
 PR 25-SEP-2000; 2000US-0235082P.

PR 25-SEP-2000; 2000US-0235134P.  
 PR 25-SEP-2000; 2000US-0235280P.  
 PR 26-SEP-2000; 2000US-0235637P.  
 PR 26-SEP-2000; 2000US-0235638P.  
 PR 27-SEP-2000; 2000US-0235711P.  
 PR 27-SEP-2000; 2000US-0235720P.  
 PR 27-SEP-2000; 2000US-0235840P.  
 PR 27-SEP-2000; 2000US-0235863P.  
 PR 28-SEP-2000; 2000US-0236028P.  
 PR 28-SEP-2000; 2000US-0236032P.  
 PR 28-SEP-2000; 2000US-0236033P.  
 PR 28-SEP-2000; 2000US-0236034P.  
 PR 28-SEP-2000; 2000US-0236109P.  
 PR 28-SEP-2000; 2000US-0236111P.  
 PR 29-SEP-2000; 2000US-0236842P.  
 PR 29-SEP-2000; 2000US-0236891P.  
 PR 02-OCT-2000; 2000US-0237172P.  
 PR 02-OCT-2000; 2000US-0237173P.  
 PR 02-OCT-2000; 2000US-0237278P.  
 PR 02-OCT-2000; 2000US-0237294P.  
 PR 02-OCT-2000; 2000US-0237295P.  
 PR 02-OCT-2000; 2000US-0237316P.  
 PR 03-OCT-2000; 2000US-0237425P.  
 PR 03-OCT-2000; 2000US-0237598P.  
 PR 03-OCT-2000; 2000US-0237604P.  
 PR 03-OCT-2000; 2000US-0237606P.  
 PR 03-OCT-2000; 2000US-0237608P.  
 PR 01-NOV-2000; 2000US-0244867P.  
 PR 01-NOV-2000; 2000US-0245084P.  
 XX (AVAIL-) AVALON PHARM.  
 PA  
 PI Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;  
 PI Sopet DR, Weaver Z;  
 XX  
 DR WPI; 2002-188264/24.  
 XX  
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical  
 PT agent to be tested for anti-neoplastic activity, and determining a change  
 PT in expression of a gene of a signature gene set.  
 XX  
 XX  
 PS Claim 1; SEQ ID NO 1680; 44pp; English.  
 XX  
 XX

CC The present invention describes a method (M1) for screening for an anti-  
 CC neoplastic agent. The method involves exposing cells to a chemical agent  
 CC to be tested for anti-neoplastic activity, determining a change in (I)  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening an  
 CC anti-neoplastic agent, and can be used for producing a product which is  
 CC the data collected with respect to the anti-neoplastic agent as a result  
 CC of M1, and the data is sufficient to convey the chemical structure and/or  
 CC properties of the agent. M1 can be used in the treatment of cancer such  
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms  
 CC tumour  
 XX  
 XX  
 SQ Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 0.00526 Length: 550  
 Score: 70.00 Matches: 14  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x ABL63343 (1-550)

OY 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
 DB |||||  
 291 GTCGTCCGACAGAGACTTAGCGCCATCTGCCCCCAACAG 332  
 RESULT 6  
 ID ABLV73813 standard; cDNA; 550 BP.  
 XX  
 AC ABLV73813;  
 XX  
 DT 08-JAN-2003 (first entry)  
 XX  
 DE Human gamma-synuclein Glu110 variant gene.  
 XX  
 KW Gamma-synuclein; human; single nucleotide polymorphism; SNP;  
 KW schizophrenia; neuroleptic; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH location/Qualifiers  
 FT 12..395  
 FT CDS  
 FT /\*tag= a  
 FT /product= "Gamma-synuclein"  
 FT /transl\_except= (pos:213..215,aa:Glu)  
 FT replace(340,T)  
 FT /\*tag= b  
 FT /standard\_name= "single nucleotide polymorphism"  
 XX  
 PN WO2002/5317-A2.  
 XX  
 PD 26-SEP-2002.  
 XX  
 PF 14-MAR-2002; 2002W0-EP002872.  
 XX  
 PR 15-MAR-2001; 2001US-0276306P.  
 XX  
 PA (NOVS ) NOVARTIS AG.  
 PA (NOVS ) NOVARTIS-ERFINDUNGEN VERM GES MBH.  
 PA (UYMA-) UNIV MARYLAND BALTIMORE.  
 XX  
 PI Roberts RC, Van Oostrum J, Voshol J, Tamminga CA;  
 PI  
 DR WPI; 2002-750574/81.  
 DR P-PSDB; ABP54932.  
 XX  
 PT Screening for compounds for treating or interfering with the onset of  
 PT Schizophrenia Spectrum Disorders, by detecting interactions of candidate  
 PT compounds with the gamma-synuclein polypeptide.  
 XX  
 PS Disclosure; Fig 1; 32pp; English.  
 XX  
 XX  
 CC The present sequence is that of cDNA encoding the Glu-110 isoform of  
 CC human gamma-synuclein. The invention relates to an isoform of gamma-  
 CC synuclein that is caused by an A/T single nucleotide polymorphism (SNP)  
 CC at position 329 of the gamma-synuclein coding sequence. This SNP causes a  
 CC glutamic acid to valine change at amino acid position 110 of gamma-  
 CC synuclein, and is associated with an increased susceptibility of  
 CC individuals to schizophrenia spectrum disorders (SSDs). This is the first  
 CC time that a genetic component of SSDs has been identified, and provides a  
 CC potential target for diagnosis and treatment of schizophrenia. Gamma-  
 CC synuclein polypeptides, especially those containing the E110V mutation,  
 CC are used in a claimed method of screening for compounds useful for the  
 CC treatment of SSDs, and gamma-synuclein expressing cells are used in a  
 CC claimed method of screening for agonist or antagonist compounds. An  
 CC oligonucleotide complementary to part of the gamma-synuclein coding  
 CC sequence is used for the discrimination of an SNP at position 329 of the  
 CC coding sequence. Gamma-synuclein polypeptides or polynucleotides are also  
 CC useful for the diagnosis of SSDs, or susceptibility to SSDs, e.g. by PCR  
 CC amplification of a polynucleotide encoding gamma-synuclein and analysis  
 CC of the occurrence of the SNP at position 329. A transgenic animal useful  
 CC for the study of SSDs is also claimed  
 XX  
 XX  
 SQ Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.00526 Length: 550  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x ABV73915 (1-550)

Qy 1 ValValArghySGluApleuArgProSerAlaProGlnGln 14  
Db 291 GTGTCGCGAAGAGAGACTTGAGGCCATCTGCCCCCAACAG 332

RESULT 7  
ABV73915  
ID ABV73915 standard; cDNA; 550 BP.

XX AC ABV73915;

XX DT 08-JAN-2003 (first entry)

XX DE Human gamma-synuclein Val110 variant gene.

XX KM Gamma-synuclein; human; single nucleotide polymorphism; SNP;  
XX KW schizophrenia; neuroleptic; mutant; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX FT CDS 12..395  
FT /\*tag= a  
FT /product= "Gamma-synuclein"  
FT /transl\_except= (pos:213..215,aa:Glu)  
FT replace(340,A)  
FT /\*tag= b  
FT /standard\_name= "Single nucleotide polymorphism"

XX PN WO200275317-A2.

XX PD 26-SEP-2002.

XX PF 14-MAR-2002; 2002MO-EP002872.

XX PR 15-MAR-2001; 2001US-0276306P.

XX PA (NOVS ) NOVARTIS AG.  
XX PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GBS MBH.  
XX PA (UYMA-) UNIV MARYLAND BALTIMORE.

XX PI Roberts RC, Van Oostrum J, Voshol J, Tamminga CA;  
XX XX  
XX XX

XX DR WPI; 2002-750574/81.  
XX DR P-PSDB; ABP54933.

XX PT Screening for compounds for treating or interfering with the onset of  
XX PT Schizophrenia Spectrum Disorders, by detecting interactions of candidate  
XX PT compounds with the gamma-synuclein polypeptide.

XX PS Disclosure; Page; 32pp; English.

XX CC The present sequence is that of cDNA encoding the Val-110 isoform of  
XX CC human gamma-synuclein. The invention relates to an isoform of gamma-  
XX CC synuclein that is caused by an A/T single nucleotide polymorphism (SNP)  
XX CC at position 329 of the gamma-synuclein coding sequence. This SNP causes a  
XX CC glutamic acid to valine change at amino acid position 110 of gamma-  
XX CC synuclein, and is associated with an increased susceptibility of  
XX CC individuals to schizophrenia spectrum disorders (SSDs). This is the first  
XX CC time that a genetic component of SSDs has been identified, and provides a  
XX CC potential target for diagnosis and treatment of schizophrenia. Gamma-  
XX CC synuclein polypeptides, especially those containing the E110V mutation,  
XX CC are used in a claimed method of screening for compounds useful for the

CC treatment of SSDs, and gamma-synuclein expressing cells are used in a  
CC claimed method of screening for agonist or antagonist compounds. An  
CC oligonucleotide complementary to part of the gamma-synuclein coding  
CC sequence is used for the discrimination of an SNP at position 329 of the  
CC coding sequence. Gamma-synuclein polypeptides or polynucleotides are also  
CC useful for the diagnosis of SSDs, or susceptibility to SSDs, e.g. by PCR  
CC amplification of a polynucleotide encoding gamma-synuclein and analysis  
CC of the occurrence of the SNP at position 329. A transgenic animal useful  
CC for the study of SSDs is also claimed. Note: The present sequence is not  
CC shown in the specification but is derived from the gamma-synuclein  
XX sequence given in Fig 1 (see ABV73915)

XX SQ Sequence 550 BP; 131 A; 145 C; 192 G; 82 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 0.00526 Length: 550  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x ABV73915 (1-550)

Qy 1 ValValArghySGluApleuArgProSerAlaProGlnGln 14  
Db 291 GTGTCGCGAAGAGAGACTTGAGGCCATCTGCCCCCAACAG 332

RESULT 8  
AAD63568  
ID AAD63568 standard; cDNA; 550 BP.

XX AC AAD63568;

XX DT 12-FEB-2004 (first entry)

XX DE Human amyloid-like protein cDNA.

XX KM Human; genetic disease; muscular dystrophy; cystic fibrosis; cytostatic;  
XX KW scientific research; gene therapy; gene; amyloid-like protein; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX FT CDS 12..395  
FT /\*tag= a  
FT /product= "Human amyloid-like protein"

XX PN US6639052-B1.

XX PD 28-OCT-2003.

XX PF 14-OCT-1999; 99US-00417540.

XX PR 30-AUG-1995; 95US-0002993P.  
XX PR 30-AUG-1996; 96US-00705771.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Moore PA;  
XX XX  
XX XX

XX DR WPI; 2003-842790/78.  
XX DR P-PSDB; ABW02024.

XX PT New isolated protein and nucleic acid molecules, useful for diagnostic  
XX PT and therapeutic purposes, e.g. for treating genetic diseases such as  
XX PT muscular dystrophy or cystic fibrosis.

XX PS Example 5; Fig 1; Opp; English.

XX CC The invention relates to isolated new isolated protein and nucleic acid  
XX CC molecules useful for diagnostic and therapeutic purposes. The invention  
XX CC is for treating genetic diseases such as muscular dystrophy or cystic

CC fibrosis, and for in vitro purposes related to scientific research,  
CC synthesis of DNA and manufacture of DNA vectors. The invention is useful  
CC in gene therapy. The present sequence is human amyloid-like protein cDNA

XX SQ Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.00526	Length:	550
Score:	70.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-017-715a-2\_COPY\_94\_107 (1-14) x AAD63568 (1-550)

OY 1 ValValArgLySGluAspLeuArgProSerAlaProGlnGln 14  
DB 291 GTGGTCCGCAAGAGACTTGAGCCATCTGCCCCCAACAG 332

RESULT 9

ADG47636  
ID ADG47636 standard; cDNA; 550 BP.

XX ADG47636;

DT 11-MAR-2004 (first entry)

XX Human amyloid like protein cDNA.

DE ss; gene; muscular dystrophy; cystic fibrosis; hypertension;

KW angina pectoris; myocardial infarction; ulcer; asthma; allergy;

KM psychosis; depression; migraine; vomiting; benign prostatic hypertrophy;

KW osteoporosis; human.

XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 12..395

XX FT /tag= a

XX FT /product= "Amyloid like protein"

XX US2003208043-A1.

XX 06-NOV-2003.

XX 04-JUN-2003; 2003US-00453478.

XX 30-AUG-1995; 95US-0002993P.

XX 30-AUG-1996; 96US-00705771.

XX 14-OCT-1999; 99US-00417540.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Moore PA, Gentz RL, Ji H, Ni J, Hu J;

XX MPI; 2003-864796/80.

XX P-PSDB; ADG47647.

XX New human polypeptides and polynucleotides, useful for diagnosing or

XX treating genetic diseases such as muscular dystrophy or cystic fibrosis,

XX hypertension, asthma, depression or osteoporosis.

XX Claim 18; SEQ ID NO 1; 56pp; English.

CC encoding human amyloid like protein.

XX SQ Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.00526	Length:	550
Score:	70.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-017-715a-2\_COPY\_94\_107 (1-14) x ADG47636 (1-550)

OY 1 ValValArgLySGluAspLeuArgProSerAlaProGlnGln 14  
DB 291 GTGGTCCGCAAGAGACTTGAGCCATCTGCCCCCAACAG 332

RESULT 10

AA29997  
ID AA29997 standard; DNA; 720 BP.

XX AA29997;

DT 06-JUL-1999 (first entry)

XX Human peryn gene.

DE Human; synuclein; peryn; diagnosis; neurodegenerative disorder; cancer;

KW breast; skin; intermediate filament damage; ss.

XX Homo sapiens.

OS Homo sapiens.

XX EP908727-A1.

XX 14-APR-1999.

XX 21-SEP-1998; 98EP-00307628.

XX 19-SEP-1997; 97GB-00019879.

XX (NEUR-) NEUROPA LTD.

XX (UYSA-) UNIV ST ANDREWS.

XX MPI; 1999-217169/19.

XX P-PSDB; AAY07271.

XX New synuclein protein (persyn) and gene, useful in assays for screening,

XX diagnosing or monitoring cancer, neurodegenerative disorders or skin

XX disorders.

XX Claim 29; Page 16-17; 39pp; English.

XX This sequence represents the gene encoding a novel human synuclein family

XX member designated persyn. The sequence is useful for screening,

XX diagnosing or monitoring cancer (especially breast or skin cancer),

XX CC neurodegenerative disorders or skin disorders and for identifying cells

XX having intermediate filament damage

XX SQ Sequence 720 BP; 173 A; 209 C; 215 G; 123 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.00711	Length:	720
Score:	70.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-017-715a-2\_COPY\_94\_107 (1-14) x AA29997 (1-720)

OY 1 ValValArgLySGluAspLeuArgProSerAlaProGlnGln 14  
DB 291 GTGGTCCGCAAGAGACTTGAGCCATCTGCCCCCAACAG 332

Db 328 GTGTGGCGAAGAGAGCTTGAGGCGACTCTGCCCCCAACAG 369

RESULT 11

ABST76519

ID ABS76519 standard; cDNA; 720 BP.

XX

XX ABS76519;

AC

XX

DT 11-DEC-2002 (first entry)

DE cDNA encoding human ovarian cancer marker OV60.

XX

XX Human; ovarian cancer; marker; cancer; familial history; brain disorder;

KM central nervous system disorder; bacterial meningitis; viral meningitis;

KM Alzheimer's disease; Parkinson's disease; cerebral edema; hydrocephalus;

KM brain herniation; inflammation; encephalitis; testicular disorder;

KM nontuberculous granulomatous orchitis; connective tissue disorder;

KM heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;

KM histological type; carcinogenic; ovarian cancer marker; gene; ss.

XX

OS Homo sapiens.

PN WO200271928-A2.

XX

PD 19-SEP-2002.

XX

PF 14-MAR-2002; 2002WO-US007826.

XX

PR 14-MAR-2001; 2001US-0276025P.

PR 14-MAR-2001; 2001US-0276026P.

PR 10-AUG-2001; 2001US-0311732P.

PR 19-SEP-2001; 2001US-0323580P.

PR 26-SEP-2001; 2001US-0324967P.

PR 26-SEP-2001; 2001US-0325102P.

PR 26-SEP-2001; 2001US-0325149P.

XX

PA (MILL-) MILLENNIUM PHARM INC.

XX

PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG,

PI Meyers RE, Morrissey WP, Olandt PJ, Sen A, Vieby PO, Mills GB,

PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;

XX

DR WPI: 2002-723277/78.

XX

PT P-PSDB; ABG96420.

XX

PT Assessing whether a patient is afflicted with ovarian cancer, useful in

PT the expression level of a cancer marker in a sample from a patient and

PT from a non cancer patient.

XX

XX

PS Disclosure; Page 411; 481pp; English.

XX

XX The present invention relates to a new method for assessing whether a

XX patient is afflicted with ovarian cancer. The method involves comparing

XX the expression level of a marker in a patient sample and the normal level

XX of expression of the marker in a control non-ovarian cancer sample, where

XX the marker is selected from 363 cancer markers described in the

XX specification. The method of the invention is useful in diagnosing or

XX characterizing cancer, in detecting the presence of cancer as early as

XX possible, and the recurrence of ovarian cancer. The method may also be of

XX particular use with patients having an enhanced risk of developing

XX ovarian cancer (e.g. patients having a familial history of ovarian

XX cancer). The cancer markers may be used in the management and treatment

XX of e.g. brain and central nervous system disorders (e.g. bacterial and

XX viral meningitis, Alzheimer's disease or Parkinson's disease), brain

XX disorders (e.g. cerebral edema, hydrocephalus or brain herniations),

XX inflammations (e.g. bacterial or viral meningitis or encephalitis),

XX testicular disorders (e.g. nontuberculous granulomatous orchitis),

XX connective tissue disorders, or heart disorders (e.g. ischemic heart

XX disease or atherosclerosis). The compositions and methods may also be

XX used in assessing the histological type of neoplasm associated with

XX ovarian cancer, monitoring the progression of ovarian cancer, determining

XX whether ovarian cancer has metastasized or is likely to metastasize,

CC selecting a composition for inhibiting ovarian cancer, assessing the

CC ovarian carcinogenic potential of a compound, or inhibiting ovarian

CC cancer or at risk of developing ovarian cancer. The present nucleic acid

CC sequence encodes one of the ovarian cancer markers described in the

CC invention

XX

SO Sequence 720 BP; 173 A; 209 C; 215 G; 123 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	720
Score:	70.00	14
Percent Similarity:	100.00%	0
Best Local Similarity:	100.00%	0
Query Match:	100.00%	0
DB:	Gaps:	0

US-09-017-715a-2\_COPY\_94\_107 (1-14) x ABS76519 (1-720)

Qy 1 ValValArgbysgluapleuArqProSerAlaProGlnGln 14

Db 328 GTGTGGCGAAGAGAGCTTGAGGCGACTCTGCCCCCAACAG 369

RESULT 12

ADE43864

ID ADE43864 standard; cDNA; 720 BP.

XX

AC ADE43864;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human SNGC cDNA, SEQ ID 469.

XX

KM Neurodegenerative disease; uPA; SNGC; IDE; KNSL1, LIPA, TNFRSF6;

KM Alzheimer's disease; neuroprotective; neurotropic; gene therapy;

KM Chromosome 10; gene; ss.

XX

OS Homo sapiens.

XX

PN WO2003054143-A2.

XX

PD 03-JUL-2003.

XX

PF 25-OCT-2002; 2002WO-US034679.

XX

PR 25-OCT-2001; 2001US-0339525P.

PR 08-NOV-2001; 2001US-0336929P.

PR 08-NOV-2001; 2001US-0338010P.

PR 09-NOV-2001; 2001US-0338363P.

PR 04-DEC-2001; 2001US-0337052P.

PR 28-MAR-2002; 2002US-0368919P.

XX

PA (NEUR-) NEUROGENETICS INC.

XX

PI (GEHO ) GEN HOSPITAL CORP.

XX

PI Becker KD, Veljoseleb G, Elliott KJ, Wang X, Tanzi RE, Bertram L,

PI Saunders AJ, Mullin KM, Sampson AJ, Blacker D;

XX

DR WPI: 2003-559131/52.

XX

PT Determining a predisposition for or the occurrence of neurodegenerative

PT disease, e.g. Alzheimer's disease by detecting in a target nucleic acid

PT the presence or absence of an allelic variant of one or more polymorphic

PT regions.

XX

XX Claim 84; Page 740; 848pp; English.

XX

XX The present invention relates to a method (M1) for determining a

XX predisposition for or the occurrence of neurodegenerative disease in a

XX subject. The method comprises detecting in a target nucleic acid obtained

XX from the subject the presence or absence of an allelic variant of one or

XX more polymorphic regions of one or more genes selected from uPA

XX (urokinase plasminogen activator), SNGC (gamma-synuclein), IDE (insulin-

XX degrading enzyme), KNSL1 (kinesin-like protein 1), LIPA (lysosomal acid

CC lysase), and TNFRSF6 (Tumour Necrosis Factor Receptor-SF6), where the  
CC presence of at least one of the allelic variant of one or more  
CC polymorphic regions is indicative of a predisposition for or the  
CC occurrence of neurodegenerative disease. The genes are all located on  
CC chromosome 10. M1 is useful for determining a predisposition for or the  
CC occurrence of, and for treating neurodegenerative disease, particularly  
CC Alzheimer's disease.

XX  
SQ Sequence 720 BP; 172 A; 205 C; 212 G; 120 T; 0 U; 11 Other;

#### Alignment Scores:

Pred. No.:	Length:	720
Score:	70.00	Matches: 14
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	10	Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x ADE43864 (1-720)

QY 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14

Db 328 GTGGTCCGACAGAGACTTGAAGCCATCTGCCCCCAACAG 369

#### RESULT 13

ADH54342  
ID ADH54342 standard; cDNA; 720 BP.

XX ADH54342;

XX 25-MAR-2004 (first entry)

XX Human SNCG gene cDNA sequence SeqID469.

XX human; neurodegenerative disease; urokinase plasminogen activator; uPA;  
KW gamma-synuclein; SNCG; insulin degrading enzyme; IDE;  
KW kinesin-like protein 1; KNSL1; lysosomal acid lipase; LIPA;  
KW tumour necrosis factor receptor SF6; TNFRSF6; Alzheimer's disease; ss.  
XX Homo sapiens.

XX US2003224380-A1.

XX 04-DEC-2003.

XX 25-OCT-2002; 2002US-00282174.

XX 25-OCT-2001; 2001US-0339525P.

XX 25-OCT-2001; 2001US-0348065P.

XX 02-NOV-2001; 2001US-0336983P.

XX 08-NOV-2001; 2001US-0336929P.

XX 09-NOV-2001; 2001US-0338610P.

XX 04-DEC-2001; 2001US-0337052P.

XX 28-MAR-2002; 2002US-0368919P.

XX (GENO) GEN HOSPITAL CORP.

XX WPI; 2004-060538/06.

XX Determining a predisposition for or the occurrence of neurodegenerative  
PT disease, particularly Alzheimer's disease, comprises determining the  
PT presence of a polymorphism in the uPA, SNCG, IDE, KNSL1, LIPA or TNFRSF6  
PT gene.

XX Claim 84; SEQ ID NO 469; 205pp; English.

CC This invention relates to a novel method of determining a predisposition  
CC for or the occurrence of neurodegenerative disease comprising detecting  
CC in a target nucleic acid obtained from the subject the presence of an

CC allelic variant of polymorphic regions of human genes selected from  
CC urokinase plasminogen activator (uPA), gamma-synuclein (SNCG), insulin  
CC degrading enzyme (IDE), kinesin-like protein 1 (KNSL1), lysosomal acid  
CC lipase (LIPA) and tumour necrosis factor receptor SF6 (TNFRSF6). The  
CC method is useful in determining the presence or predisposition to a  
CC neurodegenerative disease, particularly Alzheimer's disease. The present  
CC sequence is the cDNA sequence of the human SNCG gene which is related to  
CC the invention.

XX  
SQ Sequence 720 BP; 172 A; 205 C; 212 G; 120 T; 0 U; 11 Other;

#### Alignment Scores:

Pred. No.:	Length:	720
Score:	70.00	Matches: 14
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	12	Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x ADH54342 (1-720)

QY 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14

Db 328 GTGGTCCGACAGAGACTTGAAGCCATCTGCCCCCAACAG 369

#### RESULT 14

AA193778  
ID AA193778 standard; cDNA; 783 BP.

XX AA193778;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 13838.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation; ss.  
XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX P-PSDB; AAO13847.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 1; SEQ ID NO 13838; 139pp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AA179941-AA193841) and  
CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activity/inhibit activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and



CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pcc\_sequences

XX Sequence 783 BP, 187 A, 232 C, 237 G, 127 T, 0 U, 0 Other;

# Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 70.00	783	14	0	0	0	0
Percent Similarity: 100.00%						
Best Local Similarity: 100.00%						
Query Match: 100.00%						

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AAF21784 (1-783)

Qy 1 ValValAArgbysgluapleuArqProSeRaLaProGlnGln 14  
Db 393 GTGGTGGCAAGAGAGACTTGAGGCCATCTGCCCCCACCAG 434

## RESULT 15

AAF21784 AAF21784 standard; DNA; 796 BP.

AC AAF21784;

DT 27-MAR-2001 (first entry)

DE Human breast and ovarian cancer associated antigen gene SEQ ID 171.

XX Human: breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
XX neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
XX antidiabetic; antiinflammatory; antiviral; vulnery; anticonvulsant;  
XX antibacterial; antifungal; antiparasitic; cardiant; immune disorder;  
XX Addison's disease; allergy; autoimmune haemolytic anaemia;  
XX autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
XX multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
XX cardiovascular disorder; wound healing; neurological disease; ds.

OS Homo sapiens.

PN WO200055173-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US005881.

PR 12-MAR-1999; 99US-0124270P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

XX WPI; 2000-611515/58.

DR P-PSDB; AAB58881.

XX New human breast and ovarian cancer associated gene sequences and the  
XX polypeptides encoded by these genes, useful in the prevention, treatment  
XX and diagnosis of cancer, immune disorders, cardiovascular disorders and  
XX neurological diseases.

PS Claim 1; Page 608; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
XX proteins AAB58711 - AAB59128. The DNA and protein sequences are  
XX associated with breast and ovarian cancer. Included in the invention are  
XX sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
XX isolation and characterization of the DNA and protein sequences of the  
XX invention. The breast and ovarian cancer associated DNA, protein, agonist  
XX or antagonist sequences exhibit cytostatic; immunosuppressive; neurotropic;  
XX neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;  
XX antiinflammatory; antiviral; vulnery; anticonvulsant; antibacterial;  
XX antifungal; antiparasitic and cardiant activity. The polynucleotide and

CC protein sequences are used in the diagnosis of cancer, particularly  
CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists  
CC and antagonists may also be used in the diagnosis, prevention and treatment  
CC of immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC cardiovascular disorders such as myocardial ischemias; wound healing;  
CC neurological diseases such as cerebral anoxia and epilepsy; and  
CC infectious diseases

XX Sequence 796 BP, 195 A, 233 C, 240 G, 126 T, 0 U, 2 Other;

# Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 70.00	796	14	0	0	0	0
Percent Similarity: 100.00%						
Best Local Similarity: 100.00%						
Query Match: 100.00%						

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AAF21784 (1-796)

Qy 1 ValValAArgbysgluapleuArqProSeRaLaProGlnGln 14

Db 388 GTGGTGGCAAGAGAGACTTGAGGCCATCTGCCCCCACCAG 429

Search completed: May 4, 2005, 09:26:19  
Job time : 105.322 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 4, 2005, 09:07:55 ; Search time 795.087 Seconds  
(without alignments)  
853.206 Million cell updates/sec

Title: US-09-017-715a-2\_COPY\_94\_107  
Perfect score: 70  
Sequence: 1 VAKEDRSPAPQ 14

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODE=frame+ .p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/h/US9017715/runat\_04052005\_100744\_25608/app.query.fasta\_1.661  
-DB=Genemb1 -QFMT=faastap -SUFFIX=rge -MINMATCH=0.1 -LOOPT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US9017715\_@CGN\_1\_1\_3370\_@runat\_04052005\_100744\_25608 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEODURY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Genemb1.\*  
1: gb\_ba:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_on:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	100.0	381	9	CR541790 Homo sapi
2	70	100.0	384	9	CR541788 Homo sapi
3	70	100.0	488	9	AF411524 Homo sapi
4	70	100.0	550	6	AR412236 Sequence

5	70	100.0	550	6	AX331171	AX331171 Sequence
6	70	100.0	550	6	BD022727	BD022727 Mammary c
7	70	100.0	550	6	AF010126	AF010126 Homo sapi
8	70	100.0	704	6	CQ720882	CQ720882 Sequence
9	70	100.0	720	6	E36333	E36333 Analytical
10	70	100.0	720	6	AX004527	AX004527 Sequence
11	70	100.0	720	9	AF017256	AF017256 Homo sapi
12	70	100.0	738	11	BV177827	BV177827 segm97020
13	70	100.0	738	9	BC014098	BC014098 Homo sapi
14	53	75.7	4606	9	AF044311	AF044311 Homo sapi
15	53	75.7	5491	9	AF037207	AF037207 Homo sapi
16	53	75.7	149668	9	AC025268	AC025268 Homo sapi
17	53	75.7	175542	2	AC150062	AC150062 Gallus ga
18	53	75.7	188771	2	AC150073	AC150073 Gallus ga
19	53	75.7	190496	2	EX510354	EX510354 Homo sapi
20	49	70.0	168860	9	AL160175	AL160175 Human DNA
21	49	70.0	219205	10	AL611985	AL611985 Mouse DNA
22	48	68.6	135638	1	AF484556	AF484556 Streptomy
23	48	68.6	211075	10	AC128668	AC128668 Mus muscu
24	48	68.6	242974	2	AC103570	AC103570 Rattus no
25	47	67.1	488	10	AF518351	AF518351 Rattus no
26	47	67.1	727	6	E36334	E36334 Analytical
27	47	67.1	727	6	AX004529	AX004529 Sequence
28	47	67.1	727	10	AF017255	AF017255 Mus muscu
29	47	67.1	732	10	RNSD5YNGE	X86789 R. norvegicu
30	47	67.1	748	10	BC028508	BC028508 Mus muscu
31	47	67.1	788	5	BC075925	BC075925 Danio rer
32	47	67.1	825	6	AR383443	AR383443 Sequence
33	47	67.1	1252	3	AK114896	AK114896 Ciona int
34	47	67.1	1646	3	AK114539	AK114539 Ciona int
35	47	67.1	191707	2	CR847971	CR847971 Danio rer
36	47	67.1	247309	2	AC097837	AC097837 Rattus no
37	46	65.7	154356	10	AC131119	AC131119 Mus muscu
38	46	65.7	172435	8	OSJ000163	AL662993 Oryza sat
39	46	65.7	189771	2	AC118514	AC118514 Rattus no
40	46	65.7	246005	2	AC095834	AC095834 Rattus no
41	46	65.7	26756	1	AP000990	AP000990 Sulfolobu
42	45	64.3	465	6	CO070159	CO070159 Sequence
43	45	64.3	465	6	CO079954	CO079954 Sequence
44	45	64.3	465	6	CO136809	CO136809 Sequence
45	45	64.3	465	6	CO174616	CO174616 Sequence

## ALIGNMENTS

RESULT 1	CR541790	381 bp	MRNA	linear	PRI 29-JUN-2004
LOCUS	CR541790				
DEFINITION	Homo sapiens full open reading frame cDNA clone RZPPO834B0231D for gene SNCG, synuclein, gamma (breast cancer-specific protein 1); complete cds, without stopcodon.				
ACCESSION	CR541790				
VERSION	CR541790.1	GI:49456536			
KEYWORDS	Full ORF shuttle clone, Gateway(TM), complete cds.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Halleck A., Ebert L., Moundinya, M., Schick, M., Bisenstein, S., Korn, B., Zuo, D., Hu, Y. and Labaer, V.				
TITLE	Cloning of human full open reading frames in Gateway(TM) system entry vector (pDONR201)				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 381)				
AUTHORS	Halleck A., Ebert L., Moundinya, M., Schick, M., Bisenstein, S., Korn, B., Zuo, D., Hu, Y. and Labaer, V.				
TITLE	Direct Submision				
JOURNAL	Submitted (28-JUN-2004) RZPD Deutsches Reessourcenzentrum fuer Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany				

COMMENT RZPD: RZPD0834B0231D, ORFNo 3631  
www.rzpd.de/cgi-bin/products/c1.cgi?cloneid=RZPD0834B0231D RZPDLIB;  
Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.  
834  
www.rzpd.de/cgi-bin/products/showlib.pl.cgi/response?libNo=834  
Contact: Inge Airlart  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 100  
Fax: +49 30 32639 111  
www.rzpd.de  
This clone is available from RZPD;  
Contact RZPD (customer.service@rzpd.de) for further information.  
Clone name at Harvard Institute of Proteomics  
(www.hip.harvard.edu): RH130940.01L  
This CDS clone is part of a collection of human full ORF clones  
jointly established and verified by the Harvard Institute of  
Proteomics (HIP) and RZPD.  
This CDS has been inserted into pDONR201 via a BP Clonase(TM)  
reaction. Additional sequence has been added in front of the start  
codon: attc...AAAAA GCA GGC TCC ACC (ATG).  
The last codon is followed by the 3' att site: GACCCAGCTTTCTT.. att  
The clone is validated by full sequence check.  
Compared to the reference sequence BC014098  
we did not find any amino acid exchanges.  
Clone distribution: http://www.rzpd.de/products/orfclones/  
Location/Qualifiers  
1..381  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="RZPD0834B0231D"  
/clone\_lib="Human Full ORF Clones Gateway(TM) - RZPD"  
/lab\_host="DH5Alpha"  
/note="Vector: pDONR201, Site\_1: attP1, Site\_2: attP2"  
1..381  
/gene="SNCG"  
1..381  
/gene="SNCG"  
/codon\_start=1  
/protein\_id="CAG46589.1"  
/db\_xref="GI:49456537"  
/translation="MDVFKGFSIAKEGVGAVEKTKQGVTEAEKTEKGVVYGAKT  
KENVQSVTSVAEKTEQANAVSEAVVSVNTVARTVBEANINAVTSVAKEDLRP  
SAPQDGGASKEKEVAERASGSD"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.00392 Length: 381  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
US-09-017-715A-2\_COPY\_94\_107 (1-14) x CR541790 (1-381)  
QY 1 ValValATgLyGtLnaSpLeuAtgProSerAlaProGlnGln 14  
DB 280 GTGGTGGCAAGAGACTTGAAGCCATCTGCCCCCAACAG 321  
RESULT 2  
CR541788  
LOCUS CR541788 384 bp mRNA linear PRI 29-JUN-2004  
DEFINITION Homo sapiens full open reading frame cDNA clone RZPD0834F0930D for  
gene SNCG, synuclein, gamma (breast cancer-specific protein 1);  
complete cds, incl. stopcodon.  
ACCESSION CR541788  
VERSION CR541788.1 GI:49456532  
KEYWORDS Full ORF shuttle clone, Gateway(TM), complete cds.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 384)  
Halleck,A., Ebert,L., Mroundinya,M., Schick,M., Eisenstein,S.,  
Neubert,P., Ketrang,K., Schatten,R., Shen,B., Henze,S., Mar,W.,  
Korn,B., Zuo,D., Hu,Y. and Labaer,J.  
Cloning of human full open reading frames in Gateway(TM) system  
entry vector (pDONR201)  
TITLE Unpublished  
2 (bases 1 to 384)  
Halleck,A., Ebert,L., Mroundinya,M., Schick,M., Eisenstein,S.,  
Neubert,P., Ketrang,K., Schatten,R., Shen,B., Henze,S., Mar,W.,  
Korn,B., Zuo,D., Hu,Y. and Labaer,J.  
Direct Submission  
JOURNAL Submitted (28-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer  
Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,  
Germany  
COMMENT RZPD: RZPD0834F0930D, ORFNo 3605  
www.rzpd.de/cgi-bin/products/c1.cgi?cloneid=RZPD0834F0930D RZPDLIB;  
Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.  
834  
www.rzpd.de/cgi-bin/products/showlib.pl.cgi/response?libNo=834  
Contact: Inge Airlart  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 100  
Fax: +49 30 32639 111  
www.rzpd.de  
This clone is available from RZPD;  
Contact RZPD (customer.service@rzpd.de) for further information.  
Clone name at Harvard Institute of Proteomics  
(www.hip.harvard.edu): RH131021.01X  
This CDS clone is part of a collection of human full ORF clones  
jointly established and verified by the Harvard Institute of  
Proteomics (HIP) and RZPD.  
This CDS has been inserted into pDONR201 via a BP Clonase(TM)  
reaction. Additional sequence has been added in front of the start  
codon: attc...AAAAA GCA GGC TCC ACC (ATG).  
The stopcodon is followed by the 3' att site: GACCCAGCTTTCTT.. att  
The clone is validated by full sequence check.  
Compared to the reference sequence BC014098  
we did not find any amino acid exchanges.  
Clone distribution: http://www.rzpd.de/products/orfclones/  
Location/Qualifiers  
1..384  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="RZPD0834F0930D"  
/clone\_lib="Human Full ORF Clones Gateway(TM) - RZPD"  
/lab\_host="DH5Alpha"  
/note="Vector: pDONR201, Site\_1: attP1, Site\_2: attP2"  
1..384  
/gene="SNCG"  
1..384  
/gene="SNCG"  
/codon\_start=1  
/protein\_id="CAG46587.1"  
/db\_xref="GI:49456533"  
/translation="MDVFKGFSIAKEGVGAVEKTKQGVTEAEKTEKGVVYGAKT  
KENVQSVTSVAEKTEQANAVSEAVVSVNTVARTVBEANINAVTSVAKEDLRP  
SAPQDGGASKEKEVAERASGSD"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.00395 Length: 384  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x CR5411788 (1-384)

QY 1 ValValAlrGlySGluApLeuArgProSerAlaProGlnGln 14  
|||||  
280 GTGGTGGCAAGAGGAGCTTGAGGCCATCTGCCCCCAACAG 321

RESULT 3  
AF411524  
LOCUS AF411524 488 bp mRNA linear PRI 20-SEP-2001  
DEFINITION Homo sapiens synuclein gamma mRNA, complete cds.  
ACCESSION AF411524  
VERSION AF411524.1 GI:15705904  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE Han, C., Zhang, B., Peng, X., Yuan, J. and Qiang, B.  
TITLE Direct Submission  
JOURNAL Submitted (19-AUG-2001) Department of Biochemistry, Institute of  
Basic Medical Science, Chinese Academy of Medical Sciences, 5 Dong  
Dan San Tiao, Beijing 100005, P.R. China  
Location/Qualifiers  
1..488  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
12..395  
/note="SNGC: breast cancer-specific protein 1"  
/codon\_start=1  
/product="synuclein gamma"  
/protein\_id="AA105870.1"  
/db\_xref="GI:15705905"  
/translation="MDVFKKGFSLAKEGVDAVEKTKQVTEAEKTKEGVMYVGAKT  
KENVVSQVTSVAETKEQANAVSAVYSSVNTATKTVEAEINAIATSSGVRRDLRP  
SAQQQBEAEAKEAEVAEENQSGD"

ORIGIN

Alignment Scores:  
Pred. No.: 0.00488 Length: 488  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AF411524 (1-488)

QY 1 ValValAlrGlySGluApLeuArgProSerAlaProGlnGln 14  
|||||  
291 GTGGTGGCAAGAGGAGCTTGAGGCCATCTGCCCCCAACAG 332

RESULT 4  
AR412236  
LOCUS AR412236 550 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 1 from patent US 6639052.  
ACCESSION AR412236  
VERSION AR412236.1 GI:40167022  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.  
1 (bases 1 to 550)  
REFERENCE Moore, P. A.  
TITLE Human ADA2 polypeptides  
JOURNAL Patent: US 6639052-A 1 28-OCT-2003;  
FEATURES Location/Qualifiers  
1..550  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 0.00543 Length: 550  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AR412236 (1-550)

QY 1 ValValAlrGlySGluApLeuArgProSerAlaProGlnGln 14  
|||||  
291 GTGGTGGCAAGAGGAGCTTGAGGCCATCTGCCCCCAACAG 332

RESULT 6  
BD022727  
LOCUS BD022727 550 bp DNA linear PAT 27-AUG-2002  
DEFINITION Mammary cancer-specific gene 1.  
ACCESSION BD022727  
VERSION BD022727.1 GI:22563950  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE Ji, H. and Rosen, C. A.  
TITLE Mammary cancer-specific gene 1  
JOURNAL Patent: JP 2001509664-A 1 24-JUL-2001;  
COMMENT HUMAN GENOME SCIENCES INC  
PN JP 2001509664-A/1  
PD 24-JUL-2001

ORIGIN

Alignment Scores:  
Pred. No.: 0.00543 Length: 550  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AX331171 (1-550)

QY 1 ValValAlrGlySGluApLeuArgProSerAlaProGlnGln 14  
|||||  
291 GTGGTGGCAAGAGGAGCTTGAGGCCATCTGCCCCCAACAG 332

RESULT 5  
AX331171  
LOCUS AX331171 550 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 1680 from Patent WO0194629.  
ACCESSION AX331171  
VERSION AX331171.1 GI:18121805  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,  
Horigan, S., Soppet, D. R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
JOURNAL Patent: WO 0194629-A 1680 13-DEC-2001;  
FEATURES Location/Qualifiers  
1..550  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

PF 03-FEB-1998 JP 1996515053  
PR 03-FEB-1997 US 60/037080  
PI HONGJUN JI, CRAIG A ROSEN  
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC 10,  
C12P21/02, C12Q1/68, G01N33/574, C12N15/00, C12N5/00 CC  
Strandedness: Double;  
CC Topology: Both;  
FH Key Location/Qualifiers  
FT CDS 12..392.  
Location/Qualifiers  
1..550  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.00543 Length: 550  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x BD022727 (1-550)

QY 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
|||||  
Db 291 GTGGTGGCAGAGGAGCTTGAGGCATCTGCCCCCAACAG 332

RESULT 7  
AP010126 550 bp mRNA linear PRI 26-JUL-1997  
LOCUS Homo sapiens breast cancer-specific protein 1 (BCSG1) mRNA,  
DEFINITION complete cds.  
ACCESSION AF010126  
VERSION AF010126.1 GI:2281473  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 550)  
Ji, H., Liu, Y.E., Jia, T., Wang, M., Liu, J., Xiao, G., Joseph, B.K.,  
Rosen, C. and Shi, Y.E.  
Identification of a breast cancer-specific gene, BCSG1, by direct  
differential cDNA sequencing  
Cancer Res. 57 (4), 759-764 (1997)  
97178957  
9044857  
2 (bases 1 to 550)  
Ji, H., Liu, Y.E., Jia, T., Wang, M., Liu, J., Xiao, G., Joseph, B.K.,  
Rosen, C. and Shi, Y.E.  
Direct Submision  
Submitted (23-JUN-1997) Ped. Res., Long Island Jewish Medical  
Center, 270-05 76th Ave., New Hyde Park, NY 11040, USA  
Location/Qualifiers  
1..550  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/rnote="type="breast cancer"  
/note="cDNA highly abundant in a breast cancer library but  
not in normal library"  
1..550  
/gene="BCSG1"  
12..395  
/gene="BCSG1"  
/note="breast cancer-specific protein 1; synuclein-like;  
AD amyloid-like"  
/codon\_start=1

gene  
CDS

/product="BCSG1 protein"  
/protein\_id="BAB64109.1"  
/db\_xref="GI:2281474"  
/translation="MDVFRKGFSTIAKKVGVAVETKQGVTEAAKTEGTVYVGAKT  
KENVQSVTSVAETKEQANAVSKVSVNTVATKVEAEENIAVTSGVAKEDLRP  
SAPOQEGASKEKEVEAEAOQSGD"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.00543 Length: 550  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AF010126 (1-550)

QY 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
|||||  
Db 291 GTGGTGGCAGAGGAGCTTGAGGCATCTGCCCCCAACAG 332

RESULT 8  
CQ720882 704 bp DNA linear PAT 03-FEB-2004  
LOCUS Sequence 6816 from Patent WO02068579.  
DEFINITION CQ720882  
ACCESSION CQ720882  
VERSION CQ720882.1 GI:42281739  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kites, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
Patent: WO 02068579-A 6816 06-SEP-2002;  
PE Corporation (NY) (US)  
Location/Qualifiers  
1..704  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.00676 Length: 704  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x CQ720882 (1-704)

QY 1 ValValArgLyGluAspLeuArgProSerAlaProGlnGln 14  
|||||  
Db 327 GTGGTGGCAGAGGAGCTTGAGGCATCTGCCCCCAACAG 368

RESULT 9  
E36333 720 bp DNA linear PAT 18-JUN-2001  
LOCUS Analytical matter based on synuclein and novel synuclein protein.  
DEFINITION E36333  
ACCESSION E36333.1 GI:13022626  
VERSION E36333.1 GI:13022626  
KEYWORDS JP 1999239488-A/1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 720)

AUTHORS Andrew, S.M., Valdimia, R.B. and Arun, M.D.  
TITLE Analytical matter based on synuclein and novel synuclein protein  
JOURNAL Patent: JP 199239488-A 1 07-SEP-1999;  
COMMENT THE UNIVERSITY COURT OF THE UNIVERSITY OF ST ANDREWS, NYUROBA LTD  
OS Homo sapiens (human)  
PN JP 199239488-A/1  
PD 07-SEP-1999  
PF 21-SEP-1998 JP 1998306283  
PR 19-SEP-1997 GB 9719879.0  
PI ANDREW SMITH, MAKKARION, VALDIMIA, RUVOVICH, BUCHMAN, PI ARUN  
MILWARD DAVIS  
PC C12N15/09, A01K67/027, C12Q1/68, G01N33/53, C12N15/00 CC  
FH Key Location/Qualifiers  
FT CDS (49) (432).  
FEATURES  
source Location/Qualifiers  
1..720  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.0069 Length: 720  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-017-715a-2\_COPY\_94\_107 (1-14) x E36333 (1-720)  
Qy 1 ValValAlArglySGluAspLeuArgProSerAlaProGlnGln 14  
Db 328 GTGGTGGCAAGAGGAGCTTGAGGCCATCTGCCCAACAG 369  
RESULT 10  
LOCUS AX004527 720 bp DNA linear PAT 24-AUG-2000  
DEFINITION Sequence 1 from Patent EP0908727.  
ACCESSION AX004527  
VERSION AX004527.1 GI:9927977  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Synuclein-based assay and synuclein protein  
AUTHORS Patent: EP 0908727-A 1 14-APR-1999;  
JOURNAL NEUROPA LIMITED (GB); UNIV COURT OF THE UNIVERSITY O (GB)  
FEATURES  
source Location/Qualifiers  
1..720  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
49..432  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC04899.1"  
/db\_xref="GI:9927978"  
/translation="MDVFKKGSIAEGVAVAEKTKQGVTEAEKTKGCVWVVGAKT  
KENVVGSVTSVAEKTKQANAVSEAVVSVNTVATKTVEAEENIAVTSGVVRKEDLRP  
SAPOEGVASKEREVEAEASGSD"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.0069 Length: 720  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-017-715a-2\_COPY\_94\_107 (1-14) x AX004527 (1-720)  
Qy 1 ValValAlArglySGluAspLeuArgProSerAlaProGlnGln 14  
Db 328 GTGGTGGCAAGAGGAGCTTGAGGCCATCTGCCCAACAG 369  
RESULT 11  
LOCUS AF017256 720 bp mRNA linear PRI 23-SEP-1998  
DEFINITION Homo sapiens perSyn mRNA, complete cds.  
ACCESSION AF017256  
VERSION AF017256.1 GI:3642774  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 720)  
Ninkina, N.N., Alimova-Kost, M.V., Paterson, J.W., Delaney, L.,  
Cohen, B.B., Imreh, S., Gnuchev, N.V., Davies, A.M. and Buchman, V.L.  
Organization, expression and polymorphism of the human perSyn gene  
Hum. Mol. Genet. 7 (9), 1417-1424 (1998)  
MEDLINE 98367030  
PubMed 9700196  
REFERENCE  
2 (bases 1 to 720)  
Buchman, V.L.  
Direct Submission  
Submitted (04-AUG-1997) School of Biomedical Sciences, Univ. of St.  
Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,  
Scotland  
FEATURES  
source Location/Qualifiers  
1..720  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..48  
49..432  
/note="member of the synuclein family with a distinct  
pattern of expression"  
/codon\_start=1  
/product="perSyn"  
/protein\_id="AAC36550.1"  
/db\_xref="GI:3642775"  
/translation="MDVFKKGSIAEGVAVAEKTKQGVTEAEKTKGCVWVVGAKT  
KENVVGSVTSVAEKTKQANAVSEAVVSVNTVATKTVEAEENIAVTSGVVRKEDLRP  
SAPOEGVASKEREVEAEASGSD"  
433..720  
684..689  
706  
3' UTR  
polyA\_signal  
polyA\_site  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.0069 Length: 720  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
US-09-017-715a-2\_COPY\_94\_107 (1-14) x AF017256 (1-720)  
Qy 1 ValValAlArglySGluAspLeuArgProSerAlaProGlnGln 14  
Db 328 GTGGTGGCAAGAGGAGCTTGAGGCCATCTGCCCAACAG 369  
RESULT 12  
LOCUS BV177827 738 bp DNA linear STS 10-JUN-2004  
DEFINITION sqm7020 Human DNA (Sequenc) Homo sapiens STS genomic. sequence  
tagged site.  
ACCESSION BV177827  
VERSION BV177827.1 GI:48014020

KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE 1 (bases 1 to 738)  
Nelson, R.M., Marnellos, G., Kammerer, S., Hoyal, C.R., Shi, M.M.,  
Cantor, C.R., and Braun, A.  
Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions  
Genome Res. (2004) In press  
JOURNAL COMMENT  
Contact: Andreas Braun  
Pharmaceuticals division  
Sequenom, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 738.

FEATURES  
source  
1..738  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequenom)"  
<1..>738

ORIGIN  
STS  
Alignment Scores:  
Pred. No.: 0.00705 Length: 738  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x BVL77827 (1-738)

Qy 1 ValValaTgLySGlUaSpLeuAArgProSeAAlaProGInGln 14  
Db 330 GTGGTGGCAGAGAGACATTGAGCCATCTGCCCCCAACAG 371

RESULT 13  
BC014098  
LOCUS BC014098 758 bp mRNA linear PRI 29-JUN-2004  
DEFINITION Homo sapiens synuclein, gamma (breast cancer-specific protein 1),  
mRNA (cDNA clone MGC:20132 IMAGE:454644), complete cds.  
ACCESSION BC014098  
VERSION BC014098.2 GI:33878363  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE 1 (bases 1 to 758)  
Straussberg, R.L., Feingold, E.A., Gronow, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Dichenko, L., Marsina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Donald, M.F., Casavant, T.L.,  
Schetz, T.E., Brownstein, M.U., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abrams, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Mortley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahy, J., Helton, E., Kettelman, W., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

TITLE  
JOURNAL PubMed  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REMARK  
COMMENT  
NIH-MGC Project URL: <http://mgs.nci.nih.gov>  
On Aug 19, 2003 this sequence version replaced gi:15559464.  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,  
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stancitrop, S., Thomas, P.J., Touchman, J.W.,  
Tsurgren, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 28 Row: P Column: 20  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4507112.

FEATURES  
source  
1..758  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:20132 IMAGE:454644"  
/tissue\_type="Colon, adenocarcinoma"  
/clone\_lib="NIH MGC\_15"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
1..758  
/gene="SNCG"  
/note="synonym: SR, BCSG1"  
/db\_xref="LOCUSID:6623"  
/db\_xref="MIM:602998"  
71..454  
/gene="SNCG"  
/codon\_start=1  
/product="synuclein, gamma (breast cancer-specific protein  
1)"  
/protein\_id="AAH14098.1"  
/db\_xref="GI:15559465"  
/db\_xref="LOCUSID:6623"  
/db\_xref="MIM:602998"  
/translation="MDVRRKGSIAKEGVAGVETKQGYEAAKTEGVAVGAKT  
KENVQSVTAETKEQANVSEAVSVNTVAITVEAENIAVTSGVAKEDLRP  
SNPQGEASKEKEVAEADSGD"

ORIGIN  
Alignment Scores:  
Pred. No.: 0.00722 Length: 758  
Score: 70.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0



Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x BC014098 (1-758)

QY 1 ValValArglySGluAspLeuArgProSerAlaProGlnGln 14  
 |||||  
 Db 350 GTGCTCCGACAGAGACTTACGCCATCTCCGCCACACAG 391

RESULT 14  
 AF044311 4606 bp DNA linear PRI 29-JUL-1998  
 LOCUS Homo sapiens gamma-synuclein gene, complete cds.  
 DEFINITION AF044311  
 ACCESSION AF044311.1 GI:3347841  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 4606)  
 AUTHORS Lavedan,C., Leroy,E., Dehejia,A., Buchholtz,S., Dutra,A.,  
 Nussbaum,R.L., and Polymeropoulos,M.H.  
 TITLE Identification, localization and characterization of the human  
 gamma-synuclein gene  
 JOURNAL Hum. Genet. 103 (1), 106-112 (1998)  
 MEDLINE 98407804  
 REFERENCE 2 (bases 1 to 4606)  
 AUTHORS Lavedan,C.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-JUN-1998) NHGRI/LGDR, NIH, 49 Convent Drive,  
 Bethesda, MD 20892, USA

FEATURES  
 source  
 1..4606  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="10"  
 /map="10q23; between markers WI-5226 and AFW225YD12"  
 join(1..173,988..1029,1356..1483,3953..4024,4314..4606)  
 /product="gamma-synuclein"  
 1..52  
 join(53..173,988..1029,1356..1483,3953..4024,4314..4334)  
 /codon\_start=1  
 /product="gamma-synuclein"  
 /protein\_id="AAC27738.1"  
 /db\_xref="GI:3347842"  
 /translation="MDVFKKGFSAKEGVGAVGAEKTKQGVTEAEKTKEGVMYVGART  
 KENVVOSVTSVAEKTKEQANAVSEAVSVNTVATKTVEAEANIATVSGVVRKEDLRP  
 SAPOQSGVASKKEEVAEEAQSQSD"  
 4335..4606

ORIGIN 3'UTR

Alignment Scores:  
 Pred. No.: 38.4 Length: 4606  
 Score: 53.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 90.91% Mismatches: 0  
 Query Match: 75.71% Indels: 0  
 DB: 9 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AF044311 (1-4606)

QY 4 LysGluAspLeuArgProSerAlaProGlnGln 14  
 ::|||  
 Db 3950 CAGAGAGACTTGAGCCATCTCCGCCACACAG 3982

RESULT 15  
 AF037207 5491 bp DNA linear PRI 23-SEP-1998  
 LOCUS Homo sapiens peryn gene, complete cds.  
 DEFINITION

ACCESSION AF037207  
 VERSION AF037207.1 GI:3642902  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 5491)  
 AUTHORS Ninkina,N.N., Alimova-Kost,M.V., Paterson,J.W., Delaney,L.,  
 Cohen,B.B., Imreh,S., Gnuchev,N.V., Davies,A.W. and Buchman,V.L.  
 TITLE Organization, expression and polymorphism of the human peryn gene  
 JOURNAL Hum. Mol. Genet. 7 (9), 1417-1424 (1998)  
 MEDLINE 98367030  
 REFERENCE 2 (bases 1 to 5491)  
 AUTHORS Buchman,V.L.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-DEC-1997) Biomedical Sciences, University of St.  
 Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,  
 Scotland, UK

FEATURES  
 source  
 1..5491  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="10"  
 /map="10q23.2-q23.3"  
 join(4567..729,1544..1585,1912..2039,4515..4586,  
 4876..55169)  
 /product="peryn"  
 <567..729  
 /number=1  
 join(609..729,1544..1585,1912..2039,4515..4586,4876..4896)  
 /note="member of synuclein family; similar to the Homo  
 sapiens peryn protein encoded by sequence presented in  
 GenBank Accession Number AF017256"  
 /codon\_start=1  
 /product="peryn"  
 /protein\_id="AAC36586.1"  
 /db\_xref="GI:3642903"  
 /translation="MDVFKKGFSAKEGVGAVGAEKTKQGVTEAEKTKEGVMYVGART  
 KENVVOSVTSVAEKTKEQANAVSEAVSVNTVATKTVEAEANIATVSGVVRKEDLRP  
 SAPOQSGVASKKEEVAEEAQSQSD"  
 1544..1585  
 /number=2  
 1912..2039  
 /number=3  
 4515..4586  
 /number=4  
 4552  
 /note="polymorphism resulting in valine to glutamic acid  
 substitution"  
 /replace="a"  
 4876..5169  
 /number=5  
 /note="in human peryn mRNA clone H1 deposited in GenBank  
 Accession Number AF017256"  
 /replace="t"  
 5148..5153  
 polyA\_signal  
 polyA\_site 5169

ORIGIN

Alignment Scores:  
 Pred. No.: 44.9 Length: 5491  
 Score: 53.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 90.91% Mismatches: 0  
 Query Match: 75.71% Indels: 0  
 DB: 9 Gaps: 0

US-09-017-715A-2\_COPY\_94\_107 (1-14) x AF037207 (1-5491)

Oy 4 LysGIuAspLeuArgProSerAlaProGIn 14  
:::|||||  
Db 4512 CAGGAGACTTGAGGCGCATCTGCCCCCAACAG 4544

Search completed: May 4, 2005, 11:53:27  
Job time : 801.421 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_p2n model

Run on: May 4, 2005, 09:07:55 ; Search time 454.336 Seconds  
(without alignments)  
853.206 Million cell updates/sec

Title: US-09-017-715A-2\_COPY\_120\_127

Perfect score: 41

Sequence: 1 EBAQSGGD 8

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODE=frame+g2n.model -DEV=x1h  
-Q=/cgn2.1/USPTO.spool.h/US09017715/runat\_04052005\_100744\_25608/app\_query.fasta\_1.661  
-DB=GenBml -QFMT=fastap -SUFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pro -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09017715\_QCGN\_1\_1\_9770\_@runat\_04052005\_100744\_25608 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenBml:\*  
1: gb\_da:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_str:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

No.	Score	Query Match	Length	ID	Description
1	41	100.0	381	9	CR541790 Homo sapi
2	41	100.0	384	9	CR541788 Homo sapi
3	41	100.0	488	9	AF411524 Homo sapi
4	41	100.0	550	6	AR412236 Sequence

5	41	100.0	550	6	AX331171	AX331171 Sequence
6	41	100.0	550	6	BD022727	BD022727 Mammari c
7	41	100.0	550	6	AF010126	AF010126 Homo sapi
8	41	100.0	704	6	CQ720882	CQ720882 Sequence
9	41	100.0	720	6	E36333	E36333 Analytical
10	41	100.0	720	6	AX004527	AX004527 Sequence
11	41	100.0	720	9	AF017256	AF017256 Homo sapi
12	41	100.0	738	11	BV177827	BV177827 agm97020
13	41	100.0	758	9	BC014098	BC014098 Homo sapi
14	41	100.0	177006	2	AC102691	AC102691 Mus muscu
15	41	100.0	255619	2	AC111772	AC111772 Rattus no
16	41	100.0	286562	2	AC137025	AC137025 Rattus no
17	41	100.0	309255	2	AC133219	AC133219 Rattus no
18	38	92.7	4887	10	AY383729	AY383729 Mus muscu
19	38	92.7	10029	1	AE010406	AE010406 Methanopy
20	38	92.7	69346	2	AC134394	AC134394 Homo sapi
21	38	92.7	133288	9	AL357503	AL357503 Human DNA
22	38	92.7	182003	9	AL355987	AL355987 Human DNA
23	38	92.7	187534	10	AC132251	AC132251 Mus muscu
24	38	92.7	210950	10	AC129317	AC129317 Mus muscu
25	38	92.7	212717	10	AL929228	AL929228 Mouse DNA
26	38	92.7	246070	2	AC135737	AC135737 Mus muscu
27	37	90.2	488	10	AY518351	AY518351 Rattus no
28	37	90.2	732	10	RN6DSYNGE	X86789 R. norvegicu
29	37	90.2	1275	6	CO575643	CO575643 Sequence
30	37	90.2	3275	6	CO575642	CO575642 Sequence
31	37	90.2	22723	2	AC019971	AC019971 Drosophi
32	37	90.2	89837	2	AC091822	AC091822 Homo sapi
33	37	90.2	90633	9	AC005592	AC005592 Homo sapi
34	37	90.2	96437	9	AY601819	AY601819 Homo sapi
35	37	90.2	110000	2	LMFPCR36_10	Continuation (11 o
36	37	90.2	146000	9	AP005433	AP005433 Homo sapi
37	37	90.2	156870	2	AC149695	AC149695 Bos tauru
38	37	90.2	161117	3	AC008311	AC008311 Drosophi
39	37	90.2	162338	5	BX537341	BX537341 Zebrafish
40	37	90.2	168448	3	AC009346	AC009346 Drosophi
41	37	90.2	191734	2	AC016560	AC016560 Homo sapi
42	37	90.2	200146	2	AC073774	AC073774 Mus muscu
43	37	90.2	209523	10	AC084386	AC084386 Mus muscu
44	37	90.2	213050	1	AL646079	AL646079 Ralstonia
45	37	90.2	219405	2	AC118096	AC118096 Rattus no

## ALIGNMENTS

RESULT 1  
CR541790  
LOCUS  
DEFINITION  
Homo sapiens full open reading frame cDNA clone RZP0834B0231D for gene SNCG, synuclein, gamma (breast cancer-specific protein 1); complete cds, without stopcodon.  
CR541790  
ACCESSION  
CR541790.1 GI:49456536  
VERSION  
Full ORF shuttle clone, Gateway(TM), complete cds.  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 381)  
Halleck,A., Ebert,L., Mkundinya,M., Schick,M., Eisenstein,S., Neubert,P., Kerrang,K., Schatten,R., Shen,B., Henze,S., Mar,W., Korn,B., Zuo,D., Hu,Y. and Labaer,J.  
Cloning of human full open reading frames in Gateway(TM) system entry vector (pDONR201)  
Unpublished  
2 (bases 1 to 381)  
Halleck,A., Ebert,L., Mkundinya,M., Schick,M., Eisenstein,S., Neubert,P., Kerrang,K., Schatten,R., Shen,B., Henze,S., Mar,W., Korn,B., Zuo,D., Hu,Y. and Labaer,J.  
Direct Submission  
Submitted (28-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

COMMENT RZPD; RZPD0834B0231D, ORFNo 3631  
 www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=RZPD0834B0231D RZPDLIB;  
 Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.  
 834  
 www.rzpd.de/cgi-bin/products/showlib.pl.cgi/response?libNo=834  
 www.rzpd.de/products/orfclones/  
 Contact: Inge Axtelt  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,  
 Heubnerweg 6, D-14059 Berlin, Germany  
 Tel: +49 30 32639 100  
 Fax: +49 30 32639 111  
 www.rzpd.de  
 This clone is available from RZPD;  
 Contact RZPD (customer.service@rzpd.de) for further information.  
 Clone name at Harvard Institute of Proteomics  
 (www.hip.harvard.edu): FLH130940.01L  
 This CDS clone is part of a collection of human full ORF clones  
 jointly established and verified by the Harvard Institute of  
 Proteomics (HIP) and RZPD.  
 This CDS has been inserted without stopcodon.  
 The CDS has been inserted into pDONR201 via a BP Clonase(TM)  
 reaction. Additional sequence has been added in front of the start  
 codon: attc...AAAAA GCA GGC TCC ACC (ATG).  
 The last codon is followed by the 3' att site: GACCCAGCTTTCTT. att  
 The clone is validated by full sequence check.  
 Compared to the reference sequence BC014098  
 we did not find any amino acid exchanges.  
 Clone distribution: http://www.rzpd.de/products/orfclones/  
 Location/Qualifiers  
 1..381  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="RZPD0834B0231D"  
 /clone\_lib="Human Full ORF Clones Gateway(TM) - RZPD"  
 /lab\_host="DH5Alpna"  
 /note="Vector: pDONR201, Site\_1: attPl, Site\_2: attP2"  
 1..381  
 /gene="SNCG"  
 1..>381  
 /gene="SNCG"  
 /codon\_start=1  
 /protein\_id="CAG46589.1"  
 /db\_xref="GI:49456537"  
 /translation="MDVFKGFSIAKEGVGAVERKQGVTEAEKTKGVMYVGAKT  
 KENVQSTVSVAEKTKQKQANAYSEAVSVNTVAIKTYBEANINAVTSGVAKEDLRP  
 SAPQOEGEASKEKEVAEAOQSGDP"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 45.6 Length: 381  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatch: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-017-715A-2\_COPY\_120\_127 (1-8) x CR541790 (1-381)  
 Qy 1 Giugluaiaglnsercylgilyasp 8  
 |||||  
 Db 358 GAGAGAGCCCAAGTGGGGAGAC 381  
 RESULT 2  
 CR541788 384 bp mRNA linear PRI 29-JUN-2004  
 CR541788  
 LOCUS Homo sapiens full open reading frame cDNA clone RZPD0834F0930D for  
 DEFINITION gene SNCG, synuclein, gamma (breast cancer-specific protein 1);  
 complete cds, incl. stopcodon.  
 CR541788  
 ACCESSION CR541788.1 GI:49456532  
 VERSION Full ORF shuttle clone, Gateway(TM), complete cds.  
 KEYWORDS Homo sapiens (human)  
 SOURCE

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eubacteria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 AUTHORS 1 (bases 1 to 384)  
 Halleck,A., Ebert,L., Moundinya,M., Schick,M., Eisenstein,S.,  
 Neubert,P., Kstrang,K., Schatten,R., Shen,B., Henze,S., Mar,W.,  
 Korn,B., Zuo,D., Hu,Y. and Labaer,J.  
 Cloning of human full open reading frames in Gateway(TM) system  
 entry vector (pDONR201)  
 TITLE Unpublished  
 JOURNAL 2 (bases 1 to 384)  
 Halleck,A., Ebert,L., Moundinya,M., Schick,M., Eisenstein,S.,  
 Neubert,P., Kstrang,K., Schatten,R., Shen,B., Henze,S., Mar,W.,  
 Korn,B., Zuo,D., Hu,Y. and Labaer,J.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer  
 Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,  
 Germany  
 COMMENT RZPD; RZPD0834F0930D, ORFNo 3605  
 www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=RZPD0834F0930D RZPDLIB;  
 Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.  
 834  
 www.rzpd.de/cgi-bin/products/showlib.pl.cgi/response?libNo=834  
 www.rzpd.de/products/orfclones/  
 Contact: Inge Axtelt  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,  
 Heubnerweg 6, D-14059 Berlin, Germany  
 Tel: +49 30 32639 100  
 Fax: +49 30 32639 111  
 www.rzpd.de  
 This clone is available from RZPD;  
 Contact RZPD (customer.service@rzpd.de) for further information.  
 Clone name at Harvard Institute of Proteomics  
 (www.hip.harvard.edu): FLH131021.01X  
 This CDS clone is part of a collection of human full ORF clones  
 jointly established and verified by the Harvard Institute of  
 Proteomics (HIP) and RZPD.  
 This CDS has been inserted incl. stopcodon.  
 The CDS has been inserted into pDONR201 via a BP Clonase(TM)  
 reaction. Additional sequence has been added in front of the start  
 codon: attc...AAAAA GCA GGC TCC ACC (ATG).  
 The stopcodon is followed by the 3' att site: GACCCAGCTTTCTT. att  
 The clone is validated by full sequence check.  
 Compared to the reference sequence BC014098  
 we did not find any amino acid exchanges.  
 Clone distribution: http://www.rzpd.de/products/orfclones/  
 Location/Qualifiers  
 1..384  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="RZPD0834F0930D"  
 /clone\_lib="Human Full ORF Clones Gateway(TM) - RZPD"  
 /lab\_host="DH5Alpna"  
 /note="Vector: pDONR201, Site\_1: attPl, Site\_2: attP2"  
 1..384  
 /gene="SNCG"  
 1..384  
 /gene="SNCG"  
 /codon\_start=1  
 /protein\_id="CAG46587.1"  
 /db\_xref="GI:49456533"  
 /translation="MDVFKGFSIAKEGVGAVERKQGVTEAEKTKGVMYVGAKT  
 KENVQSTVSVAEKTKQKQANAYSEAVSVNTVAIKTYBEANINAVTSGVAKEDLRP  
 SAPQOEGEASKEKEVAEAOQSGDP"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 45.9 Length: 384  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatch: 0  
 Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x CR5411788 (1-384)

QY 1 GluGlu1aGlnSerGlyGlyasp 8  
|||  
358 GAGGAGGCCAGAGTGGGGAGAC 361

RESULT 3  
AF411524 488 bp mRNA linear PRI 20-SEP-2001  
LOCUS Homo sapiens synuclein gamma mRNA, complete cds.  
DEFINITION AF411524  
ACCESSION AF411524  
VERSION AF411524.1 GI:15705904  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 488)  
AUTHORS Han, C., Zhang, B., Peng, X., Yuan, J. and Qiang, B.  
TITLE Direct Submission  
JOURNAL Submitted (19-AUG-2001) Department of Biochemistry, Institute of  
Basic Medical Science, Chinese Academy of Medical Sciences, 5 Dong  
Dan San Tiao, Beijing 100005, P. R. China  
FEATURES  
source Location/Qualifiers  
1..488  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
12..395  
/note="SNGC: breast cancer-specific protein 1"  
/codon\_start=1  
/product="synuclein gamma"  
/protein\_id="AA05870.1"  
/db\_xref="GI:15705905"  
/translation="MDVFKKGFSLAKEGVDAVEKTKQGVTEAEKTKEGVMYVGAKT  
KENVQSVTSVAETKEQANAVSRVAVSVNTATKVEBAENIAVTSSVRRDELRLP  
SAQQQBEBAKEKEVNAEENQSGD"

ORIGIN

Alignment Scores:  
Pred. No.: 56.4 Length: 488  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AF411524 (1-488)

QY 1 GluGlu1aGlnSerGlyGlyasp 8  
|||  
Db 369 GAGGAGGCCAGAGTGGGGAGAC 392

RESULT 4  
AR412236 550 bp DNA linear PAT 18-DEC-2003  
LOCUS AR412236  
DEFINITION Sequence 1 from patent US 6639052.  
ACCESSION AR412236  
VERSION AR412236.1 GI:40167022  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 550)  
AUTHORS Moore, P. A.  
TITLE Human ADA2 polypeptides  
JOURNAL Patent: US 6639052-A 1 28-OCT-2003;  
FEATURES Location/Qualifiers  
1..550  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 62.5 Length: 550  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AR412236 (1-550)

QY 1 GluGlu1aGlnSerGlyGlyasp 8  
|||  
Db 369 GAGGAGGCCAGAGTGGGGAGAC 392

RESULT 6  
BD022727 550 bp DNA linear PAT 27-AUG-2002  
LOCUS BD022727  
DEFINITION Mammary cancer-specific gene 1.  
ACCESSION BD022727  
VERSION BD022727.1 GI:22563950  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 550)  
AUTHORS Ji, H. and Rosen, C. A.  
TITLE Mammary cancer-specific gene 1  
JOURNAL Patent: JP 2001509664-A 1 24-JUL-2001;  
COMMENT HUMAN GENOME SCIENCES INC  
PN JP 2001509664-A/1  
PD 24-JUL-2001

ORIGIN

Alignment Scores:  
Pred. No.: 62.5 Length: 550  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AX331171 (1-550)

QY 1 GluGlu1aGlnSerGlyGlyasp 8  
|||  
Db 369 GAGGAGGCCAGAGTGGGGAGAC 392

RESULT 5  
AX331171 550 bp DNA linear PAT 09-JAN-2002  
LOCUS AX331171  
DEFINITION Sequence 1680 from Patent WO0194629.  
ACCESSION AX331171  
VERSION AX331171.1 GI:18121805  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,  
Horrigan, S., Soppet, D. R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
JOURNAL Patent: WO 0194629-A 1680 13-DEC-2001;  
FEATURES Location/Qualifiers  
1..550  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

PF 03-FEB-1998 JP 1998515053  
PR 03-FEB-1997 US 60/037080  
PI HONGJUN JI, CRAIG A ROSEN  
PC  
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10,  
PC C12P21/02, C12Q1/68, G01N33/574, C12N15/00, C12N5/00 CC  
Strandedness: Double;  
CC Topology: Both;  
FH Key Location/Qualifiers  
FT CDS 12..392.  
Location/Qualifiers  
1..550  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	62.5	Length:	550
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x BD022727 (1-550)

QY 1 GIUGLUAAGInSerGlyasp 8  
|||||  
369 GAGGAGGCCAGAGTGGGAGAC 392

RESULT 7  
AP010126 550 bp mRNA linear PAT 26-JUL-1997  
LOCUS Homo sapiens breast cancer-specific protein 1 (BCSG1) mRNA,  
DEFINITION complete cds.  
ACCESSION AF010126  
VERSION AF010126.1 GI:2281473  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
REFERENCE  
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 550)  
Ji, H., Liu, Y.E., Jia, T., Wang, M., Liu, J., Xiao, G., Joseph, B.K.,  
Rosen, C. and Shi, Y.E.  
IDENTIFICATION OF A breast cancer-specific gene, BCSG1, by direct  
differential cDNA sequencing  
J Cancer Res. 57 (4), 759-764 (1997)  
JOURNAL  
MEDLINE  
PUBMED  
97178957  
9044857  
2 (bases 1 to 550)  
Ji, H., Liu, Y.E., Jia, T., Wang, M., Liu, J., Xiao, G., Joseph, B.K.,  
Rosen, C. and Shi, Y.E.  
TITLE Direct Submission  
JOURNAL Submitted (23-JUN-1997) Ped. Res., Long Island Jewish Medical  
Center, 270-05 76th Ave., New Hyde Park, NY 11040, USA  
FEATURES  
source  
1..550  
Location/Qualifiers  
1..550  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cfeature\_type="breast cancer"  
/note="cDNA highly abundant in a breast cancer library but  
not in normal library"  
1..550  
/gene="BCSG1"  
12..395  
/gene="BCSG1"  
/note="breast cancer-specific protein 1; synuclein-like;  
AD amyloid-like"  
/codon\_start=1

/product="BCSG1 protein"  
/protein\_id="AAB64109.1"  
/db\_xref="GI:2281474"  
/translation="MDVFRKGSILKKGVAVETKQGVTEAAEKTEGVVYGAKT  
KENVQSTVSVAEKTEQANAVSKAVVSVNTVAITVBEAENIAVTSGVAKEDLRP  
SAPQEGEASKEKEVAEAOGGD"

ORIGIN

Alignment Scores:

Pred. No.:	62.5	Length:	550
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AF010126 (1-550)

QY 1 GIUGLUAAGInSerGlyasp 8  
|||||  
369 GAGGAGGCCAGAGTGGGAGAC 392

RESULT 8  
CQ720882 704 bp DNA linear PAT 03-FEB-2004  
LOCUS Sequence 6816 from Patent WO02068579.  
DEFINITION CQ720882  
ACCESSION CQ720882  
VERSION CQ720882.1 GI:42281739  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
REFERENCE  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kite, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
Patent: WO 02068579-A 6816 06-SEP-2002;  
JOURNAL  
PE Corporation (NY) (US)  
FEATURES  
source  
1..704  
Location/Qualifiers  
1..704  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	77.2	Length:	704
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x CQ720882 (1-704)

QY 1 GIUGLUAAGInSerGlyasp 8  
|||||  
405 GAGGAGGCCAGAGTGGGAGAC 428

RESULT 9  
E36333 720 bp DNA linear PAT 18-JUN-2001  
LOCUS Analytical matter based on synuclein and novel synuclein protein.  
DEFINITION E36333  
ACCESSION E36333.1 GI:13022626  
VERSION E36333.1  
KEYWORDS JP 1999239488-A/1.  
SOURCE Homo sapiens (human)  
ORGANISM  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 720)

**AUTHORS** Andrew, S.M., Valdimira, R.B. and Arun, M.D.  
**TITLE** Analytical matter based on synuclein and novel synuclein protein  
**JOURNAL** Patent: JP 1999239488-A 1 07-SEP-1999;  
 THE UNIVERSITY COURT OF THE UNIVERSITY OF ST ANDREWS, NUTROBA LTD

**COMMENT** OS Homo sapiens (human)  
 PN JP 1999239488-A/1  
 PD 07-SEP-1999  
 PF 21-SEP-1998 JP 1998306283  
 PR 19-SEP-1997 GB 9719879.0  
 PI ANDREW SMITH MAKARION, VALDIMIRA RIVOICHI BUCHIMAN, PI ARUN  
 MILMAD DAVIS  
 PC C12N15/09, A01K67/027, C12Q1/68, G01N33/53, C12N15/00 CC  
 FH Key Location/Qualifiers  
 FT CDS (49) . (432).

**FEATURES** Location/Qualifiers  
 source 1..720  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

# ORIGIN

## Alignment Scores:

Pred. No.:	78.7	Length:	720
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x E36333 (1-720)

**QY** 1 GluGlualaglnSerglygIyasp 8  
 Db 406 GAGGAGGCCAGAGTGGGGAGAC 429

**RESULT 10**  
 AX004527 720 bp DNA linear PAT 24-AUG-2000  
 LOCUS Sequence 1 from Patent EP0908727.  
 AX004527  
 ACCESSION AX004527.1 GI:9927977  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

**REFERENCE** 1 Synuclein-based assay and synuclein protein  
**AUTHORS** Patent: EP 0908727-A 1 14-APR-1999;  
**TITLE** NEUROPA LIMITED (GB); UNIV COURT OF THE UNIVERSITY O (GB)  
**JOURNAL** Location/Qualifiers

**FEATURES** Location/Qualifiers  
 source 1..720  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 49..432  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC04899.1"  
 /db\_xref="GI:9927978"  
 /translation="MDVFKKGSIAKEGVGAVETKQGVTEAEKTKEGVMYVGAKT  
 KENVVSVSVAKETKEQANVSEAVVSVNTVATKVEAEINIAVTSGVVRKEDLRP  
 SAPOQGVASKEKEVEAEBOAGSGD"

## CDS

**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 78.7 Length: 720  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AX004527 (1-720)

**QY** 1 GluGlualaglnSerglygIyasp 8  
 Db 406 GAGGAGGCCAGAGTGGGGAGAC 429

**RESULT 11**  
 AF017256 720 bp mRNA linear PRI 23-SEP-1998  
 LOCUS Homo sapiens peryn mRNA, complete cds.  
 AF017256  
 ACCESSION AF017256.1 GI:3642774  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

**AUTHORS** Minkina, N.N., Alimova-Kost, M.V., Paterson, J.W., Delaney, L.,  
 Cohen, B.B., Imreh, S., Gnuchey, N.V., Davies, A.M., and Buchman, V.L.  
**TITLE** Organization, expression and polymorphism of the human peryn gene  
**JOURNAL** Hum. Mol. Genet. 7 (9), 1417-1424 (1998)  
**REFERENCE** 98367030  
**PUBMED** 9700196  
**AUTHORS** 2 (bases 1 to 720)  
**TITLE** Buchman, V.L.  
**JOURNAL** Direct Submission  
 Submitted (04-AUG-1997) School of Biomedical Sciences, Univ. of St.  
 Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,  
 Scotland

**FEATURES** Location/Qualifiers  
 source 1..720  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 1..48  
 49..432  
 /note="member of the synuclein family with a distinct  
 pattern of expression"  
 /codon\_start=1  
 /product="peryn"  
 /protein\_id="AAC36550.1"  
 /db\_xref="GI:3642775"  
 /translation="MDVFKKGSIAKEGVGAVETKQGVTEAEKTKEGVMYVGAKT  
 KENVVSVSVAKETKEQANVSEAVVSVNTVATKVEAEINIAVTSGVVRKEDLRP  
 SAPOQGVASKEKEVEAEBOAGSGD"

## CDS

**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 78.7 Length: 720  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AF017256 (1-720)

**QY** 1 GluGlualaglnSerglygIyasp 8  
 Db 406 GAGGAGGCCAGAGTGGGGAGAC 429

**RESULT 12**  
 BV177827 738 bp DNA linear STS 10-JUN-2004  
 LOCUS sqm97020 Human DNA (Sequenc) Homo sapiens STS genomic, sequence  
 DEFINITION tagged site.  
 BV177827  
 ACCESSION BV177827.1 GI:48014020  
 VERSION





Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x BC014098 (1-758)

Db	428	GAGGAGGCCACAGATGTGGGAGAC	451
RESULT 14	<p>1 GIJGU1AAGInserG yGlyasp 8</p> <p>     </p>		
LOCUS	AC102691		
DEFINITION	Mus musculus chromosome 10 clone RP24-273A2 map 10. *** SEQUENCING IN PROGRESS ***; 6 unordered pieces.		
ACCESSION	AC102691		
VERSION	AC102691.3 GI:45861072		
SOURCE	HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVERTIN.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	<p>Mus musculus</p> <p>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.</p> <p>1 Birren, B., Nusbbaum, C. and Lander, E.</p> <p>Mus musculus chromosome 10, clone RP24-273A2</p> <p>Unpublished</p> <p>2 (bases 1 to 177006)</p>		
AUTHORS	<p>Birren, B., Linton, L., Nusbbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Chepel, Y., Colangelo, M., Collins, S., Collinmore, A., Cook, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kelle, C., Larocque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, U., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPherson, R., Meldrim, J., Menusz, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Notbu, C., Notman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riebeck, K., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stenge-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, W.</p>		
TITLE	Direct Submission		
JOURNAL	Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 177006)		
AUTHORS	<p>Birren, B., Nusbbaum, C., Lander, E., Abouelkell, A., Allen, N., Anderson, N., Anderson, S., Archchik, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chepel, Y., Collinmore, A., Cook, A., Cooke, P., Corum, B., DeRellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafe, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, U., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menusz, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Notbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stenge-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkatarman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,</p>		

[illegible]

REFERENCE 1 (bases 1 to 255619)  
Rattus.  
Muzny D, Marie, Melker M, Lee, Abramson S, Adams C, Alder J, Allen C, Allen H, Alsbrooks S, Amin A, Angiano D, Anyalebech V, Ayagi A, Ayodeji M, Baca E, Baden H, Baldwin D, Bandaranaike D, Barber M, Barnstead M, Benahmed F, Bismail K, Blair J, Blankenburg K, Blyth P, Brown M, Bryant N, Burch P, Burch P, Burrell K, Calderon E, Cardenas V, Carter K, Cavazos I, Cesar H, Center A, Chacko J, Chavez D, Chen G, Chen R, Chen Y, Chen Z, Chu J, Cleveland C, Cockrell R, Cox C, Coyle M, Cree A, D'Amico L, Davila M, Davis C, Davy-Carroll L, De Anda C, Dederich D, Delgado O, Denison S, Deramo C, Ding Y, Dinh H, Divya K, Draper H, Dugan-Rocha S, Dunn A, Durbin K, Duval B, Eaves K, Egan A, Escotto M, Eugene C, Evans C, A, Falls T, Fan G, Fernandez S, Finley M, Flagg N, Forbes L, Foster M, Foster P, Fraser C, Gabisi A, Gatta R, Garcia A, Garner T, Garza M, Gebregorgis E, Geer K, Gill R, Grady M, Guerra M, Guevara W, Gunaratne P, Haaland W, Hamil C, Hamilton C, Hamilton K, Harvey Y, Havlak P, Hawes A, Henderson N, Hernandez J, Hernandez R, Hines S, Hladun S, L, Hodgson A, Hognes M, Hollins B, Howells S, Hulyk S, Hume J, Idlebird D, Jackson A, Jackson L, Jacob L, Jiang H, Johnson B, Johnson R, Jolyet A, Karpathy S, Kelly S, Kelly S, Khan Z, King L, Kovar C, Kowis C, Kraft C, Lebow H, Levan J, Lewis S, Li Z, Liu J, Liu J, Liu W, Liu Y, London P, Longacre S, Lopez J, Lounsbury L, Louleed H, Lozano R, Lu X, Ma J, Maheshwari M, Mahindaratne M, Mahmoud R, Malloy K, Mangum A, Mangum B, Mapua P, Martin K, Martin R, Martinez E, Mawhinney S, McLeod M, McNeill T, Z, Meenen E, Milosavljevic A, Miner G, Minja E, Montemayor J, Moore S, Morgan M, Morris K, Morris S, Mundaasa M, Murphy M, Nait L, Nankervis C, Neal D, Newton N, Nguyen N, Norris S, Nwankweli O, Okunolu G, Olarnunsgoon A, Pal S, Parks K, Pasternak S, Paul H, Perez A, Perez L, Pfankuch C, Plouffe F, Poindexter A, Popovic D, Primus E, Pu L, L, Pucio M, Quirio J, Rachlin E, Reeves K, Regier M, A, Reigh R, Reilly B, Reilly M, Ren Y, Reuter M, Richards S, Riggs F, Rivers C, Rodkey T, Rojas A, Rose M, Rose R, Ruiz S, J, Sanders W, Savary G, Scherer S, Scott G, Shatsman S, Shen H, Shetty J, Shvartsbeyn A, Sison I, Sitter C, D, Smajd, Sneed A, Sodergren E, Song X, Z, Sorrelle R, Sosa D, Steidle M, Strong R, Sutton A, Svatek A, Taber P, Taylor C, Taylor T, Thomas N, Thomas S, Tingey A, Trejos Z, Usmani K, Valsar R, Vera V, Villaseana D, Waldron L, Walker B, Wang J, Williams G, Wang S, Warren J, Warren R, Wei X, White F, Wright D, Wright R, Wu J, Yakub S, Yen U, Yoon L, Yoon V, Yu F, Zhang U, Zhou J, Zhou X, Zhao S, Dunn D, von Niederhausern A, Weiss R, Smith D, R, Holt R, A, Smith H, O, Weinstock G, and Gibbs R, A.

TITLE JOURNAL  
REFERENCE 2 (bases 1 to 255619)  
Worley K, C.  
Direct Submission  
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

TITLE JOURNAL  
REFERENCE 3 (bases 1 to 255619)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT  
On Nov 15, 2002 this sequence version replaced gi:3365002.  
The sequence in this assembly is a combination of BAC based reads and whole shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GOBO  
Center clone name: CH230-169C7  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 212404 bases at least Q40  
Consensus quality: 215959 bases at least Q30  
Consensus quality: 217955 bases at least Q20  
Estimated insert size: 218715; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a "working draft" sequence. It currently consists of 6 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 7834: contig of 7834 bp in length  
\* 7835 7934: gap of unknown length  
\* 7935 18981: contig of 11047 bp in length  
\* 18982 19081: gap of unknown length  
\* 19082 62574: contig of 4393 bp in length  
\* 62575 62674: gap of unknown length  
\* 62675 65974: contig of 3300 bp in length  
\* 65975 66074: gap of unknown length  
\* 66075 254373: contig of 188299 bp in length  
\* 254374 254473: gap of unknown length  
\* 254474 255619: contig of 1146 bp in length.

FEATURES  
source 1..255619  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-169C7"  
complement(6735..7529)  
/note="clone\_boundary"  
clone\_end:5p6  
site:  
end sequence: BH336366"  
7935..9481  
misc\_feature  
/note="wgs\_contig"  
19082..20748  
misc\_feature  
/note="wgs\_contig"  
23484..25134  
misc\_feature  
/note="wgs\_contig"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.21e+04 Length: 255619  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-09-017-715a-2\_COPY\_120\_127 (1-8) x AC111772 (1-255619)

QY 1 GIUGIUAAGInSergIvGIvYasp 8  
|||||

Thu May 5 15:10:43 2005

us-09-017-715a-2\_copy\_120\_127.rge

Page 9

Db 223025 GAGAGCAGAGAGTGGAGGGGAC 223048

Search completed: May 4, 2005, 11:54:02  
Job time : 489.669 secs

---

**THIS PAGE LEFT BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 4, 2005, 09:07:55 ; Search time 7212.58 Seconds  
(without alignments)  
853.206 Million cell updates/sec

Title: US-09-017-715A-2

Perfect score: 610  
Sequence: 1 MDVFKKGFSIAKKGVGAVG.....EGGASKEKEVAHEHAGSGD 127

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool.h/US09017715/runat.04052005.100744.25608/app.query.fasta.1.661  
-DB=GenEmbl -QFMT=fastap -SUFFIX=ixe -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USERS=US09017715.OCGN.1.1.3970.0runat.04052005.100744.25608 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MAIN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_hvg: \*  
3: gb\_in: \*  
4: gb\_cm: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	610	100.0	550	6	AR412236 Sequence
2	610	100.0	550	6	AX331171 Sequence
3	610	100.0	550	6	BD022727 Mammary c
4	610	100.0	550	9	AF010126 Homo sapi

5	602	98.7	381	9	CR541790 Homo sapi
6	602	98.7	384	9	CR541788 Homo sapi
7	602	98.7	704	6	CQ720882 Sequence
8	602	98.7	758	9	BC014098 Homo sapi
9	595	97.5	488	9	AF411524 Homo sapi
10	595	97.5	720	6	E36333 Analytical
11	595	97.5	720	6	AX004527 Sequence
12	595	97.5	720	6	AF017256 Homo sapi
13	532.5	87.3	738	11	BV177827 sqm97020
14	523	85.7	677	4	AF219257 Bos tauru
15	497	81.5	727	6	E36334 Analytical
16	497	81.5	727	6	AX004529 Sequence
17	497	81.5	727	6	AF017255 Mus muscu
18	497	81.5	748	10	BC028508 Mus muscu
19	476	78.0	488	10	AY518351 Rattus no
20	470	77.0	732	10	RNSDSYNGE
21	438.5	71.9	873	5	AF253513 Gallus ga
22	412	67.5	1120	5	CR762140 Xenopus t
23	399.5	65.5	1046	5	BC054269 Xenopus l
24	388.5	63.7	1040	5	AY055119 Xenopus l
25	388.5	63.7	1180	5	BC072217 Xenopus l
26	316	51.8	695	10	S73007 synuclein S
27	316	51.8	1017	10	AY550005 Rattus no
28	316	51.8	1017	10	AY550006 Rattus no
29	316	51.8	1018	10	AF007758 Rattus no
30	316	51.8	1047	10	BC046764 Mus muscu
31	316	51.8	1124	10	AF033261 Mus muscu
32	316	51.8	1181	10	AF179273 synuclein S
33	315	51.6	695	10	S73008 synuclein S
34	314.5	51.6	1234	5	SEISYNELFI
35	311.5	51.1	1182	5	AF253512
36	311.5	51.1	1190	5	EX936116
37	311	51.0	1233	5	BC054200 Xenopus l
38	310	50.8	543	10	AF044672 Mus muscu
39	308	50.5	549	10	DI7764 Rattus sp.
40	307.5	50.4	420	9	CR541653 Homo sapi
41	307.5	50.4	423	6	AX662310 Sequence
42	307.5	50.4	423	9	CR457058 Homo sapi
43	307.5	50.4	423	12	BT007765 Synthetic
44	307.5	50.4	425	9	AY049786 Homo sapi
45	307.5	50.4	546	6	CQ728889 Sequence

ALIGNMENTS

RESULT 1	AR412236	Sequence 1 from patent US 6639052.	550 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	AR412236					
DEFINITION	Sequence 1 from patent US 6639052.					
ACCESSION	AR412236					
VERSION	AR412236.1	GI:40167022				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 550)					
AUTHORS	Moore,P.A.					
TITLE	Human ADA2 polypeptides					
JOURNAL	Patent: US 6639052-A 1 28-OCT-2003;					
FEATURES	Location/Qualifiers					
source	1..550					
ORIGIN	/organism="unknown"					
	/mol_type="genomic DNA"					

Alignment Scores:	7.25e-44	length:	550
Pred. No.:	610.00	Matches:	127
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	6	Gaps:	0
DB:			

US-09-017-715A-2 (1-127) x AR412236 (1-550)

QY 1 MetaspvAlPheLyserGlyPheSerIleAlaLyseGlyValValaGlu 20  
DB 12 ATGAGATGTTTCAAGAGGCGCTTCTCCATCGCCAAAGGCGGTGGGTGGGAGAA 71  
QY 21 LysThrLyseGlnGlyValThrGluAlaGluLyseThrLyseGluGlyValMetTyVal 40  
DB 72 AAGACCAAGCAGGGGGTGAAGAGAGAGCTGAGAAACCAAGAGGGGGTCAATGATGTG 131  
QY 41 G1yAlaLyseThrLyseGluAsnValValGlnSerValThrSerValAlaGluLyseThrLyse 60  
DB 132 GGAGCCCAAGACCAAGAGATGTTGTACAGAGCGTACCTCGGGTGGCCAGAGCAAG 191  
QY 61 GluGlnAlaAsnAlaValSerLyseValValSerSerValAsnThrValAlaThrLyse 80  
DB 192 GAGCAGGCGCAAGCGCTGTGAGCAAGGCTGTGAGAGGCTCAACTGTGGCCACCAAG 251  
QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValArgLyseGluAsn 100  
DB 252 ACCGTGAGAGAGCGGAGAACATCGCGTCACTCGGGGTGGTGCAGAGAGAGACTTG 311  
QY 101 ArgProSerAlaProGlnGlnGluGlyGluAlaSerLyseGluLyseGluValAlaGlu 120  
DB 312 AGGCCATCTGCCCGCCCAAGAGAGGCGTGAAGCATCCAAAGAGAGAGAGTGGCAGAG 371  
QY 121 GluAlaGlnSerGlyGlyAsp 127  
DB 372 GAGGCCCAAGATGGGGAGAC 392

RESULT 2  
AX331171  
LOCUS AX331171 550 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 1680 from Patent WO0194629.  
ACCESSION AX331171  
VERSION AX331171.1 GI:18121805  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,  
Hortigan, S., Soppet, D. R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 1680 13-DEC-2001;  
FEATURES Avalon Pharmaceuticals (US)  
source 1. .550 Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

## Alignment Scores:

Pred. No.: 7, 25e-44 Length: 550  
Score: 610.00 Matches: 127  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-017-715A-2 (1-127) x AX331171 (1-550)

QY 1 MetaspvAlPheLyserGlyPheSerIleAlaLyseGlyValValaGlu 20  
DB 12 ATGAGATGTTTCAAGAGGCGCTTCTCCATCGCCAAAGGCGGTGGGTGGGAGAA 71  
QY 21 LysThrLyseGlnGlyValThrGluAlaGluLyseThrLyseGluGlyValMetTyVal 40  
DB 72 AAGACCAAGCAGGGGGTGAAGAGAGAGCTGAGAAACCAAGAGGGGGTCAATGATGTG 131

QY 41 G1yAlaLyseThrLyseGluAsnValValGlnSerValThrSerValAlaGluLyseThrLyse 60  
DB 132 GGAGCCCAAGACCAAGAGATGTTGTACAGAGCGTACCTCGGTGGCCAGAGAGCAAG 191  
QY 61 GluGlnAlaAsnAlaValSerLyseValValSerSerValAsnThrValAlaThrLyse 80  
DB 192 GAGCAGGCGCAAGCGCTGTGAGCAAGGCTGTGAGACAGGTCACACTGTGGCCACCAAG 251  
QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValArgLyseGluAsn 100  
DB 252 ACCGTGAGAGAGCGGAGAACATCGCGTCACTCGGGGTGGTGCAGAGAGAGACTTG 311  
QY 101 ArgProSerAlaProGlnGlnGluGlyGluAlaSerLyseGluLyseGluValAlaGlu 120  
DB 312 AGGCCATCTGCCCGCCCAAGAGAGGCGTGAAGCATCCAAAGAGAGAGAGTGGCAGAG 371  
QY 121 GluAlaGlnSerGlyGlyAsp 127  
DB 372 GAGGCCCAAGATGGGGAGAC 392

RESULT 3  
BD022727  
LOCUS BD022727 550 bp DNA linear PAT 27-AUG-2002  
DEFINITION Mammary cancer-specific gene 1.  
ACCESSION BD022727  
VERSION BD022727.1 GI:22563950  
KEYWORDS JP 2001509664-A/1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Ji, H. and Rosen, C. A.  
TITLE Mammary cancer-specific gene 1  
JOURNAL Patent: JP 2001509664-A 1 24-JUL-2001;  
COMMENT HUMAN GENOME SCIENCES INC  
PN JP 2001509664-A/1  
PD 24-JUL-2001  
PF 03-FEB-1998 JP 1998515053  
PR 03-FEB-1997 US 60/037080  
PI HONGJUN JI, CRAIG A ROSEN

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10,  
PC C12P21/02, C12Q1/68, G01N33/574, C12N15/00, C12N5/00 CC  
Strandedness: Double;  
CC Topology: Both; Location/Qualifiers  
FH Key 12. .392.  
FT CDS Location/Qualifiers  
source 1. .550  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

## ORIGIN

## Alignment Scores:

Pred. No.: 7, 25e-44 Length: 550  
Score: 610.00 Matches: 127  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-017-715A-2 (1-127) x BD022727 (1-550)

QY 1 MetaspvAlPheLyserGlyPheSerIleAlaLyseGlyValValaGlu 20  
DB 12 ATGAGATGTTTCAAGAGGCGCTTCTCCATCGCCAAAGGCGGTGGGTGGGAGAA 71  
QY 21 LysThrLyseGlnGlyValThrGluAlaGluLyseThrLyseGluGlyValMetTyVal 40  
DB 72 AAGACCAAGCAGGGGGTGAAGAGAGAGCTGAGAAACCAAGAGGGGGTCAATGATGTG 131

QY	60
QY	41
Db	132
QY	61
Db	192
QY	81
Db	252
QY	101
Db	312
QY	121
Db	372
RESULT 4	
LOCUS	AF010126
DEFINITION	Homo sapiens breast cancer-specific protein 1 (BCSG1) mRNA, complete cds.
VERSION	AF010126
KEYWORDS	AF010126.1 GI:2281473
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (baaes 1 to 550) Ji, H., Liu, Y. E., Jia, T., Wang, M., Liu, J., Xiao, G., Joseph, B. K., Rosen, C. and Shi, Y. E.
TITLE	Identification of a breast cancer-specific gene, BCSG1, by direct differential cDNA sequencing
JOURNAL	Cancer Res. 57 (4), 759-764 (1997)
MEDLINE	97178957
PUBMED	9044857
REFERENCE	2 (baaes 1 to 550) Ji, H., Liu, Y. E., Jia, T., Wang, M., Liu, J., Xiao, G., Joseph, B. K., Rosen, C. and Shi, Y. E.
AUTHORS	Direct Submission
TITLE	Submitted (22-JUN-1997) Ped. Res., Long Island Jewish Medical Center, 270-05 76th Ave., New Hyde Park, NY 11040, USA
JOURNAL	Location/Qualifiers
FEATURES	1..550
source	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /cfeature="breast cancer" /note="cDNA highly abundant in a breast cancer library but not in normal library"
gene	1..550
CDS	/gene="BCSG1" 12..395 /gene="BCSG1" /note="breast cancer-specific protein 1; synclestin-like; AD amyloid-like" /codon_start=1 /product="BCSG1 protein" /protein_id="AA064109.1" /db_xref="GI:2281474" /translation="MDVFKKFSIAKGVGVGAEKQGVTEAAEKTEKGVVYGATTTKENVVQSTVAEKTEQANAVSKAVVSVNTVATKTVEEENIAVTSVGRKEDLRPSAPDQGEASKEKEBEVAEASQSGD"
ORIGIN	
Alignment Scores:	7.25e-44
Pred. No.:	Length: 550
Score:	610.00
	Matches: 127

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0
US-09-017-715A-2 (1-127) x AF010126 (1-550)			
QY	1 MetaspValPheIyVblySGIyPhSeSerIleAlaIylySGIyValValGIyAlaValGIu 20		
Db	12 ATGATGATGTTTTCAGAAAGGGGCTTCTCCATCCCAAGAAAGGGCGTGCGGTGGAA 71		
QY	21 LyethrlySGInGIyValIThrGIuAlaAlaGIuLyethrlySGIyValMetTyVal 40		
Db	72 AAGACCAAGCGGGGGTGAAGGACGACTGAAGAACCAAGGAGGGGGTATGATGTG 131		
QY	41 GIyAlaIyThrlySGIuAsnValaIGInSeSerValThSeSerValaIGIuLySGThlys 60		
Db	132 GGACCCAAGACCAAGGAGAAATGTTGTACAGAGCCTGACCTAGTGGCCGAAACCAAG 191		
QY	61 GIuGIuAlaAsnAlaValaISerlySAlaValaISerSeSerValaAnthrValaIThrlys 80		
Db	192 GAGCAGGCCCAAGCCCGTGACCAAGCGTGTGGTAGCAGCCGCAACACTGTGGCCACCAAG 251		
QY	81 ThrValaGIuGIuAlaGIuAsnIleAlaValaThSeSerGIyValaIyGIySGIuAspIeu 100		
Db	252 ACCGTGAGGAGGGGGGAAACATGCGGTCCCTCCGGGGTGGTGGCCGAAAGAGACTTG 311		
QY	101 ArgProSeSerAlaPcGInGIuGIuGIyAlaIaSerlySGIuLySGIuValaIaGIu 120		
Db	312 AGGCATCTGCGCCCCCAACAGAGAGGTGAGCAATCCAAAGAAAGAGAGTGGCAGAG 371		
QY	121 GIuAlaGIuSerGIyGIyAsp 127		
Db	372 GAGGCCCAAGAGTGGGGAGAGAC 392		
RESULT 5			
CR541790	381 bp	mRNA	linear
LOCUS			
DEFINITION	Homo sapiens full open reading frame cDNA clone RZPD0834B0231D for gene SNGC, synuclein, gamma (breast cancer-specific protein 1); complete cds, without stopcodon.		
ACCESSION	CR541790		
VERSION	CR541790.1	GI:49456536	
KEYWORDS	Full ORF shuttle clone, Gateway(TM), complete cds.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 381) Halleck,A., Ebert,L., Mkundinya,M., Schick,M., Bisenstein,S., Neubert,P., Kstrang,K., Schatten,R., Shen,B., Henze,S., Mar,W., Korn,B., Zuo,D., Hu,Y. and Labaer,J.		
TITLE	Cloning of human full open reading frames in Gateway(TM) system entry vector (pDONR201)		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 381) Halleck,A., Ebert,L., Mkundinya,M., Schick,M., Bisenstein,S., Neubert,P., Kstrang,K., Schatten,R., Shen,B., Henze,S., Mar,W., Korn,B., Zuo,D., Hu,Y. and Labaer,J.		
AUTHORS	Direct Submission Submitted (28-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany		
TITLE	RZPD: RZPD0834B0231D, ORFno 3631		
JOURNAL	www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=RZPD0834B0231D RZPDLIB: Human Full ORF Clones Gateway(TM) - RZPD (kan-resistc.) RZPD LIB No. 834		
COMMENT	www.rzpd.de/cgi-bin/products/showlib.pl.cgi?response?libno=834 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH, Heubenerweg 6, D-14059 Berlin, Germany Tel: +49 30 32639 100		

Fax: +49 30 32639 111

www.rzpd.de

This clone is available from RZPD:

Contact RZPD (customer.service@rzpd.de) for further information.

Clone name at Harvard Institute of Proteomics

(www.hip.harvard.edu): FLH130940.01L

This clone is part of a collection of human full ORF clones

jointly established and verified by the Harvard Institute of

Proteomics (HIP) and RZPD.

This CDS has been cloned without stopcodon.

The CDS has been inserted into pDONR201 via a BP Clonase(TM)

reaction. Additional sequence has been added in front of the start

codon: att. AAAA GCA GGC TCC ACC (ATG).

The last codon is followed by the 3' att site: GACCAGCTTCTT. att

Compared to the reference sequence BC014098

We did not find any amino acid exchanges.

Clone distribution: http://www.rzpd.de/products/orfclones/.

#### FEATURES

##### source

1. 381

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="RZPD0834B0231D"

/clone\_lib="Human Full ORF Clones Gateway(TM) - RZPD"

/lab\_host="DH5Alpha"

/note="Vector: pDONR201, Site\_1: attPl, Site\_2: attp2"

1. 381

/gene="SNCG"

1. 381

/gene="SNCG"

/codon\_start=1

/protein\_id="CAG46589.1"

/db\_xref="GI:49456537"

/translation="MDVFRKFSIAKEGVGAVETKQGVTEAKTEKGVNYGAKT

KENVQSVTAETKEQANAVSEAVSVNTVATKYVEAENINAVTSGVAREDLRP

SAPOQEGEASKEKEVEAEBOAGSDP"

#### ORIGIN

##### Alignment Scores:

Pred. No.: 2,38e-43 Length: 381  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 9 Gaps: 0

US-09-017-715A-2 (1-127) x CR541790 (1-381)

QY 1 MetAspValaPhelyselysGlyPheSerIleAlaIyLySgIyValValaGlyAlaValaGlu 20

Db 1 ATGAGNGCTTTCAGAGAGGCTTCTCCATCCCAAGAGGCGCTGTGGTGGCGTGGAA 60

QY 21 LysThrLySgIngIyValThrGluAlaIaGluLyThrLySgIngIyValMetIyVal 40

Db 61 AAGACCAACAGAGGGGTGACGAGACGCTGAGAAACCAAGAGGCGGTCAATGATGTG 120

QY 41 GlyAlaLyThrLySgIngIyValaGlnInservValThrServAlaIaGluLySgThrLy 60

Db 121 GGAGCCCAAGACCAAGAGATGTTGACAGAGGTAACCTCACTGGCCGAGAAACCAAG 180

QY 61 GluGlnAlaAsnAlaValSerLySgAlaValaIseSerValaAsnThrValaIaThrLy 80

Db 181 GAGCAGGCGCAAGCCCTGACGAGGCTGTGGTACAGCCTCAACCTGTGGCCACCAAG 240

QY 81 ThrValaGluGluAlaGluAsnIleAlaValaThrServIyValaIaIaGlySgIyAspLeu 100

Db 241 ACCGTGAGAGAGCGGAGAACATCGCGTCACTCCGGGGTGGTGCAGAGAGGACTTG 300

QY 101 ArgProSerAlaProGlnIngIyGluGluAlaSerLySgIyLySgIyGluValaIaGlu 120

Db 301 AGGCCATCTGCCCCCAAGAGAGGCTGAGGATCCAAAGAGAAAGAGAAAGTGGCAGAG 360

QY 121 GualaGlnSerGlyIyAsp 127  
Db 361 GAGGCCAGAGTGGGAGAGAC 381

#### RESULT 6

##### LOCUS

##### DEFINITION

CR541788 384 bp mRNA linear PRI 29-JUN-2004

Homo sapiens full open reading frame CDNA clone RZPD0834F0930D for

gene SNCG, synuclein, gamma (breast cancer-specific protein 1);

complete cds, incl. stopcodon.

CR541788

CR541788.1 GI:49456532

Full ORF shuttle clone, Gateway(TM), complete cds.

Homo sapiens (human)

#### ORGANISM

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### REFERENCE



/codon\_start=1  
/protein\_id="CAG46587.1"  
/db\_xref="GI:49456533"  
/translation="MDVFKKGFSTAKGVGVGAVETKQGTVEAEKTEKGVVYVCAAT  
KENVQSVTSVAEKTQEVANAVSEAVSVNTATKTVEEANIATVSGVVRKEDLRP  
SAPOQGEASKEKEVEAEASQSGD"

## ORIGIN

## Alignment Scores:

Pred. No.: 2.4e-43 Length: 384  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 9 Gaps: 0

US-09-017-715a-2 (1-127) x CR541788 (1-384)

Qy 1 MetAspValPheIyLysGlyPheSerIleAlaIyLysGlyValAlaGlyAlaValGlu 20  
Db 1 ATGATGTCTTCAAGAGGCTTCTTCATCCCAAGAGGCGTGTGGTCCGATGAA 60  
Qy 21 LysThrLysGlnGlyValThrGluAlaAlaGluLysThrLysGlnGlyValMetTyVal 40  
Db 61 AAGACCAAGCAGGGGGTGAAGGAGCAAGCTGAGAAAGCAAGAGGGGTGATGATGTG 120  
Qy 41 G1yAlaIyLysThrLysGluAsnValAlaGlnSerValThrSerValAlaGluLysThrLys 60  
Db 121 GAGACCAAGCAGGAGAAATGTGTACAGAGCTGACCTCAGTCCGCGAGAACCAAG 180  
Qy 61 G1uGlnAlaAsnAlaValSerLysAlaValAlaSerSerValAsnThrValAlaThrLys 80  
Db 181 GAGCAGGCCAAGCGCGTGAAGAGAGCTGTGTGACAGCTCAACACTGTGGCCACCAAG 240  
Qy 81 ThrValG1uGluAlaGluAsnIleAlaValThrSerGlyValAlaArgLysGluAspLeu 100  
Db 241 ACCGTGAGAGAGGGCGAAGACATCGCGGTACCTCCGGGGTGGTCCGAGAGGACTTG 300  
Qy 101 ArgProSerAlaProGlnGlnGlnGlyGluAlaSerLysGlyLysGluGluValAlaGlu 120  
Db 301 AGGCATCTGCCCGCCCAAGAGAGGTGAGGATCCAAAGAAAGAGAGAGTGGCAGAG 360  
Qy 121 G1uAlaGlnSerGlyLysAsp 127  
Db 361 GAGGCCAGAGTGGCGAGAG 381

RESULT 7  
LOCUS CQ720882 704 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 6816 from Patent WO02068579.  
ACCESSION CQ720882  
VERSION CQ720882.1 GI:42281739

KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.  
Kits, such as nucleic acid arrays, comprising a majority of  
humans or transcripts, for detecting expression and other uses  
thereof

JOURNAL Patent: WO 02068579-A 6816 06-SEP-2002;  
PE Corporation (NY) (US)

FEATURES  
source Location/Qualifiers  
1..704

ORIGIN  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

Alignment Scores: 4.59e-43 Length: 704  
Pred. No.: 4.59e-43 Length: 704

Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 6 Gaps: 0

US-09-017-715a-2 (1-127) x CQ720882 (1-704)

Qy 1 MetAspValPheIyLysGlyPheSerIleAlaIyLysGlyValAlaGlyAlaValGlu 20  
Db 48 ATGATGTCTTCAAGAGGCTTCTTCATCCCAAGAGGCGTGTGGTCCGATGAA 107  
Qy 21 LysThrLysGlnGlyValThrGluAlaAlaGluLysThrLysGlnGlyValMetTyVal 40  
Db 108 AAGACCAAGCAGGGGGTGAAGGAGCAAGCTGAGAAAGCAAGAGGGGTGATGATGTG 167  
Qy 41 G1yAlaIyLysThrLysGluAsnValAlaGlnSerValThrSerValAlaGluLysThrLys 60  
Db 168 GAGCAGGCCAAGCGCGTGAAGAGAGCTGTGTGACAGCTCAACACTGTGGCCACCAAG 227  
Qy 61 G1uGlnAlaAsnAlaValSerLysAlaValAlaSerSerValAsnThrValAlaThrLys 80  
Db 228 GAGCAGGCCAAGCGCGTGAAGAGAGCTGTGTGACAGCTCAACACTGTGGCCACCAAG 287  
Qy 81 ThrValG1uGluAlaGluAsnIleAlaValThrSerGlyValAlaArgLysGluAspLeu 100  
Db 288 ACCGTGAGAGAGGGCGAAGACATCGCGGTACCTCCGGGGTGGTCCGAGAGGACTTG 347  
Qy 101 ArgProSerAlaProGlnGlnGlnGlyGluAlaSerLysGlyLysGluGluValAlaGlu 120  
Db 348 AGGCATCTGCCCGCCCAAGAGAGGTGAGGATCCAAAGAAAGAGAGAGTGGCAGAG 407  
Qy 121 G1uAlaGlnSerGlyLysAsp 127  
Db 408 GAGGCCAGAGTGGCGAGAG 428

RESULT 8  
LOCUS BC014098 758 bp mRNA linear PRI 29-JUN-2004  
DEFINITION Homo sapiens synuclein, gamma (breast cancer-specific protein 1),  
mRNA (cDNA clone MGC:20132 IMAGE:454644), complete cds.  
ACCESSION BC014098  
VERSION BC014098.2 GI:33878363  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
Strausberg,R.L., Pingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,  
Altschul,S.F., Zeeberg,B., Bueltow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
Schertzer,T.E., Brownstein,M.U., Uedini,T.B., Toshiyuki,S.,  
Carninci,P., Prange,C., Kana,S.S., Loquelliano,N.A., Peters,G.J.,  
Adamsom,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,  
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huiyk,S.W.,  
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Fahy,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Boutard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Butterfield,Y.S., Krzywinski,M.I., Skalski,U., Smalins,D.E.,  
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.

JOURNAL Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16699-16903 (2002)  
REFERENCE 12477932  
AUTHORS Strausberg,R.

TITLE Direct Submission  
JOURNAL Submitted (10-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Aug 19, 2003 this sequence version replaced gi:15559464.  
COMMENT Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantipop, S., Thomas, P.D., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

FEATURES  
source  
1. 758  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:20132 IMAGE:454644"  
/tissue\_type="Colon, adenocarcinoma"  
/clone\_id="NIH\_MGC\_15"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
1. 758  
/gene="SNCG"  
/note="synonyms: SR, BCSG1"  
/db\_xref="LOCUSID:6623"  
/db\_xref="MIM:602998"  
71. 454  
/gene="SNCG"  
/codon\_start=1  
/product="synuclein, gamma (breast cancer-specific protein 1)"  
/protein\_id="AAH14098.1"  
/db\_xref="GI:15559465"  
/db\_xref="LOCUSID:6623"  
/db\_xref="MIM:602998"  
/translation="MDVFKGFSIAKEGVGAVERTKQGVTEAEKTEGMYVGAKT KENVQSVTAETKEQANAVSEAVSVNTVATKVEAEENIAVTSGVAKEDLRP SAPQOGEASKEKEVAERAOEGCD"

## ORIGIN

## Alignment Scores:

Pred. No.: 4,97e-43 Length: 758  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 9 Gaps: 0

US-09-017-715A-2 (1-127) x BC014098 (1-758)

QY 1 MetAaPvAlPhelYelYsGlyPheSerIleAlaLYsGlyValValaGlu 20  
DB 71 ATGGATGCTTCAAGAGGGCTTCTCCATCGCCAAAGAGGGCGTGCGTGGTGGGAA 130

QY 21 LysThrLYSGInGlyValThrGluAlaGluLYsThrLYSGInGlyValMetTYrVal 40  
DB 111 MAGACCAAGCAGGGGTGATCGAAGCACTGAGAGACCAAGAGGGGTCTGATG 190  
QY 41 GJYAlaLYsThrLYSGInGlyValaSerValaGlnSerValaThrSerValaGluLYsThrLYs 60  
DB 131 GGAGCCAAAGACCAAGAGATGTTGTACAGACCGTACCTCAGTGGCCAGAACAA 250  
QY 61 GJUGInAlaSerAlaValaSerValaValaSerValaAspThrValaLaThrLYs 80  
DB 251 GAGCAGGCCAAAGCGGTGAGCGAGCTGTGTGACACGCTCAACACTGTGGCCAA 310  
QY 81 ThrValaGluAlaGluAlaSerAlaValaThrSerGlyValaValaArgLYsGluAspLeu 100  
DB 311 ACCGTGAGAGAGCGGAGAACATCGCGTCACTCCGGGGTGTCCGACAGAGACTTG 370  
QY 101 ArgProSerAlaProGlnGlnGlnGlyGluAlaSerLYsGlyGluGluValaGlu 120  
DB 371 AGCCATCTGCCCGCCCAAGAGAGGTGAGCATCCAAAGAAAGAGAGTGGCAG 430  
QY 121 GJUAaGlnSerGlyValaAsp 127  
DB 431 GAGGCCAGAGTGGGAGAC 451

RESULT 9  
AF411524 488 bp mRNA linear PRI 20-SEP-2001  
LOCUS AF411524  
DEFINITION Homo sapiens synuclein gamma mRNA, complete cds.  
ACCESSION AF411524  
VERSION AF411524.1 GI:15705904  
KEYWORDS

SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS Han, C., Zhang, B., Peng, X., Yuan, J. and Qiang, B.  
TITLE Direct Submission  
JOURNAL Submitted (19-AUG-2001) Department of Biochemistry, Institute of Basic Medical Science, Chinese Academy of Medical Sciences, 5 Dong Dan San Tiao, Beijing 100005, P.R. China

FEATURES  
source  
1. 488  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
12. 395  
/note="SNCG; breast cancer-specific protein 1"  
/codon\_start=1  
/product="synuclein gamma"  
/protein\_id="AAI05870.1"  
/db\_xref="GI:15705905"  
/translation="MDVFKGFSIAKEGVDAVEKTKQGVTEAEKTEGMYVGAKT KENVQSVTAETKEQANAVSEAVSVNTVATKVEAEENIAVTSGVAKEDLRP SAPQOGEASKEKEVAERAOEGCD"

## CDS

ORIGIN  
Alignment Scores:  
Pred. No.: 1,24e-42 Length: 488  
Score: 595.00 Matches: 124  
Percent Similarity: 99.21% Conservative: 2  
Best Local Similarity: 97.64% Mismatches: 1  
Query Match: 97.54% Indels: 0  
DB: 9 Gaps: 0

US-09-017-715A-2 (1-127) x AF411524 (1-488)

QY 1 MetAaPvAlPhelYelYsGlyPheSerIleAlaLYsGlyValValaGlu 20  
DB 12 ATGGATGCTTCAAGAGGGCTTCTCCATCGCCAAAGAGGGCGTGATGATCGGAGAA 71  
QY 21 LysThrLYSGInGlyValThrGluAlaGluLYsThrLYSGInGlyValMetTYrVal 40

Df		72	AAGACCAAGCAGGGGGTGACGGAACAGCTTGAAAGACCAGAGGGGGTCATGTATGTG	131
OY		41	GjyalalyThrLygLuasnValajInserValThrSerValAlagluThrlyS	60
Df		132	GGAGCCAAAGACCAAGAGATGTGTGAACAAGCGTACTCATGTGGCGAAGAACAAG	191
OY		61	GlUGlUlnlaaenValalSerlyValAlaValalSerSerValAsnThValAlathrlYS	80
Df		192	GAGCAGGGCCAACGCCGTGAGACGCCTGTGGTGGACAGCGCTCACACTGTGGCCACCAAG	251
OY		81	ThrValglUGluAlaGlUasnIleAlavalThrserglValValArxLysglUaspLeu	100
Df		252	ACCGTGGAGAGAGCCGAGAACATCCCGGTCACTCCGGGGTGTCTCCGAAAGAGGACTTG	311
OY		101	ArgProSerAlaProlngInglUglUalaserlySGluLYsgLUglUValAlaglu	120
Df		312	AGGCCATCTGCCCCCCCAAGAGAGGTGAGCATCCAAAGAAAGAGAGTGCACAG	371
OY		121	GLUALagInserGLyLYasp	127
Df		372	GAGGCCAGAGTGGGGGAGAC	392
RESULT 10				
LOCUS	E36333	720 bp	DNA	linear PAT 18-JUN-2001
DEFINITION	Analytical matter based on synuclein and novel synuclein protein.			
ACCESSION	E36333			
VERSION	E36333.1 GI:13022626			
KEYWORDS	JP 1999239488-A/1.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
REFERENCE	1 (bases 1 to 720)			
AUTHORS	Andrew,S.M., Valdimia,R.B. and Arun,M.D.			
TITLE	Analytical matter based on synuclein and novel synuclein protein			
JOURNAL	Patent: JP 1999239488-A 1 07-SEP-1999;			
	THE UNIVERSITY COURT OF THE UNIVERSITY OF ST ANDREWS,NYUROBA LTD			
COMMENT	OS Homo sapiens (human)			
	PN JP 1999239488-A/1			
	PD 07-SEP-1999			
	PF 21-SEP-1998 JP 1998306283			
	PR 19-SEP-1997 GB 9719879.0			
	PI ANDREW SMITH MAKKARION,VALDIMIA RUVOVICH I BOCHIMAN, PI ARUN			
	MIAMARD DAVIS			
	PC C12N15/09,A01K67/027,C12J01/68,G01N33/53,C12N15/00 CC			
	FH Key Location/Qualifiers			
	FT CDS Location/Qualifiers (49)..(432).			
FEATURES				
source	1..720 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"			
ORIGIN				
Alignment Scores:				
Pred. No.:	1.88e-42	Length:	720	
Score:	595.00	Matches:	124	
Percent Similarity:	99.21%	Conservative:	2	
Best Local Similarity:	97.64%	Mismatch:	1	
Query Match:	97.54%	Indels:	0	
Df		Gaps:	0	
US-09-017-715A-2 (1-127) x E36333 (1-720)				
OY		1	MetaspValPhelyLySGlyPheserlleAlalySLyglYvalValGlyAlaValajcu	20
Df		49	ATGAGTGTCTTCAAGAGAGGGCTTCCATCCCAAGAGAGGGCGTGGTGGCGGTGGA	108
OY		21	LysThrLySLyngInglYalThrGUlnAlaAglyUsThrLySGlyglYValMetTYTal	40
Df		109	AAGACCAAGCAGGGGGTGACGGAACAGCTTGAAAGACCAGAGGGGGTCATGTATGTG	168

Oy		41	G Y A L A T S T R V S G U A S N V A I G I S E R V A I T H S E R V A I A G U L V S T R V S	60
Dd		169	GGAAGCCAAAGCAACGAGGAATGTCTTCACAGCGTGACCTCAGTCGCCGAGAACACCAAG	228
Oy		61	G U G U N A A S N A V A S E T V S A V A V S E S E R V A A S N V A I A T H V S	80
Dd		229	GAGCAGGCACCAACGGCGTGAGCGAAGCTGTGGTGTAGCAGCGTCACACTGTGTGCCACCAG	288
Oy		81	T h r v a l g u g u a g u a e n i l e a l v a t h s e r g v a l v a a g v g u a p l e u	100
Dd		289	ACCTGTGAGGAGGGGGGAGAACTGCGCGTCACTCCGGGTGTGCCCAAGAGGACTTG	348
Oy		101	A r g P o s e r a A P C O G I N G I N G U G I V A I S E T V S G U L V S G U G U V A I A G U	120
Dd		349	AGGCACTCTGCCCCCAACAGAGAGGCTGTGCATCCAAAGAAAGAAAGAAAGTAGTGCAGAG	408
Oy		121	G U A A G I N S E R G I Y A E P	127
Dd		409	GAGGCCAGAGTGGGGGAGAC	429
<b>RESULT 11</b>				
AX004527			720 bp	DNA linear PAT 24-AUG-2000
LOCUS		AX004527		
DEFINITION		Sequence 1 from Patent EP0908727.		
ACCESSION		AX004527		
VERSION		AX004527.1		GI:9927977
KEYWORDS				
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens		
REFERENCE		1		
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
JOURNAL		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
FEATURES				
source		synuclein-based assay and synuclein protein Patent: EP 0908727-A 1 14-Apr-1999; NEUROPA LIMITED (GB); UNIV COURT OF THE UNIVERSITY O (GB) Location/Qualifiers 1..720		
CDS		/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 49..432 /note="unnamed protein product"		
ORIGIN		/codon_start=1 /protein_id="CAC04899.1" /db_xref="GI:9927978" /translation="MDVFKKGFISIAEGVVAGVEKTKQGVTEAEAKTKEGVMYVGAKT KENNVSVTSVAEKTRQANAASEAVWSVNTATTVBEAENIAVTSGVRREDLRP SAPOEGVASKEKEVEAAEASGSD"		
Alignment Scores:				
Pred. No.:		1.88e-42	Length:	720
Score:		595.00	Matches:	124
Percent Similarity:		99.21%	Conservative:	2
Best Local Similarity:		97.64%	Mismatches:	1
Query Match:		97.54%	Indels:	0
DB:		6	Gaps:	0
US-09-017-715A-2 (1-127) x AX004527 (1-720)				
Oy		1	M e f A s p V A P h e t V S L V S G I P h e S e r I e a l V S L V S G I V A I G I V A I G U	20
Dd		49	A T G A T G T C T T C A G A A G G C T T C T C A T G C C A A G A G G G C G T G T G G T G G A	108
Oy		21	L V E T H V S G I N J V A I T H G U A A A G U L V S T R V S G U G I V A I M E T V A I	40
Dd		109	A A G C C A A G C A G G G T G A C G A A G C G T G A C A A G C C A A G A G G G G C G T A T G T	168
Oy		41	G Y A L A T S T R V S G U A S N V A I G I S E R V A I T H S E R V A I A G U L V S T R V S	60
Dd		169	GGAAGCCAAAGCAACGAGGAATGTCTTCACAGCGTGACCTCAGTCGCCGAGAACACCAAG	228

QY 61 GUGUUAAsnAlaValSerLySaAlaValSerSerValAsnThrValAlaThrLyS 80  
 |||...|||  
 Db 229 GAGCAGGCCAACGCGGTGAGCAGCGTGTGAGCAGCGTCAACCTGTGCCACCAAG 288  
 QY 81 ThrValGUGUUAAGUAsnIleAlaValThrSerGlyValValArgLySGLuAspleu 100  
 |||...|||  
 Db 289 ACCGTGAGAGAGCGGAGAACATCGCGTCACTCCGGGTGTGGCAGAGAGACTTG 348  
 QY 101 ArgProSerAlaProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120  
 |||...|||  
 Db 349 AGCCCATCTGCCCCCACAAGAGAGGCTGTGCATCCAAAGAGAAAGAGAGAGAGAG 408  
 QY 121 GUUAGlnSerGlyGlyAsp 127  
 |||...|||  
 Db 409 GAGGCCAGAGTGGGGAGAC 429

RESULT 12  
 AF017256 720 bp mRNA linear PRI 23-SEP-1998  
 LOCUS Homo sapiens persyn mRNA, complete cds.  
 DEFINITION AF017256  
 ACCESSION AF017256.1 GI:3642774  
 VERSION AF017256.1 GI:3642774  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 720)  
 AUTHORS Ninkina,N.N., Alimova-Kost,M.V., Paterson,J.W., Delaney,L.,  
 Cohen,B.B., Imreh,S., Gnuchev,N.V., Davies,A.M. and Buchman,V.L.  
 TITLE Organization, expression and polymorphism of the human persyn gene  
 JOURNAL Hum. Mol. Genet. 7 (9), 1417-1424 (1998)  
 MEDLINE 98367030  
 PUBMED 9700196  
 REFERENCE 2 (bases 1 to 720)  
 AUTHORS Buchman,V.L.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-AUG-1997) School of Biomedical sciences, Univ. of St.  
 Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,  
 Scotland

FEATURES  
 source Location/Qualifiers  
 1..720  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 1..48  
 49..432  
 /note="member of the synuclein family with a distinct  
 pattern of expression"  
 /codon\_start=1  
 /product="persyn"  
 /protein\_id="AAC36550.1"  
 /db\_xref="GI:3642775"  
 /translation="MDVFKGPSIAKEGVAGVEKTKGVTAAEKTEGVYVGAKT  
 KENVVOSTVSAETKEQANAVSEAVSVNTVATVTVEANINAVTSVVKEDLRP  
 SAPQEGVSAKEKEVEAEAGSGD"  
 433..720  
 polyA\_signal  
 polyA\_site  
 684..689  
 706

Alignment Scores:  
 Pred. No.: 1,886-42 Length: 720  
 Score: 595.00 Matches: 124  
 Percent Similarity: 99.21% Conservative: 2  
 Best Local Similarity: 97.64% Mismatches: 1  
 Query Match: 97.54% Indels: 0  
 Gaps: 0  
 DB: 9  
 US-09-017-715A-2 (1-127) x AF017256 (1-720)

QY 1 MetAspValPheLySlyGlyPheSerIleAlaLySlyGlyValValGlyAlaValGln 20

Db 49 ANGATGTCCTTCAAGAGGGCTTCCATCCGACAGAGAGGGGTGTGGTCCGATCGAA 108  
 |||...|||  
 QY 21 LysThrLySGLnGlyValThrGluAlaValGlnGlnGlnGlnGlnGlnGlnGln 40  
 |||...|||  
 Db 109 AAGACCAAGCAGGGGTGTGCGGACAGCTGTGAAGACCAAGAGGGGTCTGTATGTG 168  
 QY 41 GYAlaLyThrLySGLuAsnValValGlnSerValThrSerValAlaGlnLyThrLyS 60  
 |||...|||  
 Db 169 GAGCCAGACCAAGAGAGATTTTACAGACCTGACTCTGATGCTCCAGAGACCAAG 228  
 QY 61 GUGUUAAsnAlaValSerLySaAlaValSerSerValAsnThrValAlaThrLyS 80  
 |||...|||  
 Db 229 GAGCAGGCCAACGCGGTGAGCAGCGTGTGAGCAGCGTCAACCTGTGCCACCAAG 288  
 QY 81 ThrValGUGUUAAGUAsnIleAlaValThrSerGlyValValArgLySGLuAspleu 100  
 |||...|||  
 Db 289 ACCGTGAGAGAGCGGAGAACATCGCGTCACTCCGGGTGTGGCAGAGAGACTTG 348  
 QY 101 ArgProSerAlaProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120  
 |||...|||  
 Db 349 AGCCCATCTGCCCCCACAAGAGAGGCTGTGCATCCAAAGAGAAAGAGAGAGAG 408  
 QY 121 GUUAGlnSerGlyGlyAsp 127  
 |||...|||  
 Db 409 GAGGCCAGAGTGGGGAGAC 429

RESULT 13  
 BV177827 738 bp DNA linear STS 10-JUN-2004  
 LOCUS sqm97020 Human DNA (Sequenom) Homo sapiens STS genomic, sequence  
 DEFINITION BV177827  
 ACCESSION BV177827.1 GI:48014020  
 VERSION BV177827.1 GI:48014020  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 738)  
 AUTHORS Nelson,R.M., Warrnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,  
 Cantor,C.R. and Braun,A.  
 TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
 JOURNAL Regions  
 COMMENT Genome Res. (2004) In press  
 CONTACT: Andreas Braun  
 Pharmaceuticals division  
 Sequenom, Inc.  
 3595 John Hopkins Court, San Diego, CA 92121, USA  
 Tel: 18582029018  
 Fax: 18582029020  
 Email: abraun@sequenom.com  
 Primer A: No primer sequence submitted  
 Primer B: No primer sequence submitted  
 STS size: 738.

FEATURES  
 source Location/Qualifiers  
 1..738  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Human DNA (Sequenom)"  
 <1..>738

Alignment Scores:  
 Pred. No.: 4,486-37 Length: 738  
 Score: 532.50 Matches: 120  
 Percent Similarity: 96.06% Conservative: 2  
 Best Local Similarity: 94.49% Mismatches: 5  
 Query Match: 87.30% Indels: 2  
 DB: 11  
 Gaps: 0

US-09-017-715A-2 (1-127) x BV177827 (1-738)

QY 1 MetaspValPheIysIysGlyPheSerIleAlaIysIysGlyValValGlyAlaValGlu 20  
Db 53 ATGGATGCTTCAAGAAAGGGCTCTCCATCCCAAGAGGGCGTGGTGGCGGTGAA 112  
QY 21 LysThrIysGlnGlyValThrGluAlaAlaGluIysThrIysGluIysValMetTyVal 40  
Db 113 TAGCCAAAGCAGTGTGTTACGAAAGCACTGAGAACCAAGAGGGGTG-ATGTATGT- 170  
QY 41 GlyAlaIysThrIysGluAenValValGlnSerValThrSerValAlaGluIysThrIys 60  
Db 171 GGAGCCAAAGCAAGAGAAATGTTGTACAGAGCGTGAAGCTCAGTGGCCGAGAACCAAG 230  
QY 61 GlnGlnAlaAenAlaValSerIysAlaValIserSerValAenThrValAlaThrIys 80  
Db 231 GAGAGGCAAGCAAGCGCTGAGAGAGCTGTGTGAGACCGTCAACACTGTGCCACCAAG 290  
QY 81 ThrValGlnGluAlaGluAenIleAlaValThrSerGlyValAlaArgIysGluAspLeu 100  
Db 291 ACCGTGAGAGAGCGGAGAAACATGCGGTCACTCCGGGGTGTGCCAGAGAGACTTG 350  
QY 101 ArgProSerAlaProGlnGlnGluIysGluIysSerIysGluIysGluValAlaGlu 120  
Db 351 AGGCACTGTGCCCCCAAGAGAGGTGGGCTATCCAAAGAAAGAGAGTGGCAGAG 410  
QY 121 GlnAlaGlnSerGlyIysAsp 127  
Db 411 GAGGCCAGAGTGGGGAGAC 431

## RESULT 14

AF219257 677 bp mRNA linear MAM 09-FEB-2000  
LOCUS Bos taurus synoretin mRNA, complete cds.  
DEFINITION AF219257  
ACCESSION AF219257.1 GI:6942173  
VERSION  
KEYWORDS

SOURCE  
ORGANISM Bos taurus (cow)

REFERENCE  
AUTHORS Surguchov, A., Surgucheva, I., Solesio, E. and Baehr, W.  
TITLE Synoretin-A new protein belonging to the synuclein family  
JOURNAL Mol. Cell. Neurosci. 13 (2), 95-103 (1999)  
MEDLINE 99210388  
PUBMED 10192768

REFERENCE  
AUTHORS Surguchov, A., Surgucheva, I., Baehr, W. and Solesio, E.  
TITLE Direct Submission  
JOURNAL Submitted (27-DEC-1999) Ophthalmology, Washington University School of Medicine in St. Louis, 660 South Euclid, St. Louis, MO 63108, USA

FEATURES  
source location/Qualifiers

CDS 1..677  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
1..384  
/note="member of gamma-subfamily of synuclein; similar to perSyn; expressed in retina and brain"  
/codon\_start=1  
/product="synoretin"  
/protein\_id="AAF2342.1"  
/db\_xref="GI:6942174"  
/translation="MDVFKKGFSLAKGVGAVETKQGVTEAAKTKEGVMYGAATKEGVSVTSVAETKEQANASVAVSVNTATKVEVENIAVTSVYHKAALQPVRSQEDBAKAEQVAEETKSGD"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.67e-36 Length: 677

Score: 523.00 Matches: 109  
Percent Similarity: 91.34% Conservative: 7  
Best Local Similarity: 85.83% Mismatches: 11  
Query Match: 85.74% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2 (1-127) x AF219257 (1-677)

QY 1 MetaspValPheIysIysGlyPheSerIleAlaIysIysGlyValValGlyAlaValGlu 20  
Db 1 ATGGACGCTTCAAGAAAGGGCTCTCCATTCGCAAGAGGGGTGGTGGCGGTGAA 60  
QY 21 LysThrIysGlnGlyValThrGluAlaAlaGluIysThrIysGluIysValMetTyVal 40  
Db 61 AAGCCAAAGCAAGAGGGGTGTTGTCCAGTGTGAGCTTCAAGAGGGGTGAGAACCAAG 120  
QY 41 GlyAlaIysThrIysGluAenValValGlnSerValThrSerValAlaGluIysThrIys 60  
Db 121 GGAGCTAAAGCAAGAGGGGTGTTGTCCAGTGTGAGCTTCAAGTGGCTGAGAACCAAG 180  
QY 61 GlnGlnAlaAenAlaValSerIysAlaValIserSerValAenThrValAlaThrIys 80  
Db 181 GAGCAGGCCAAAGCGCTGAGAGAGCGCTGTCTCCAGTGTCAACACTGTGCCACCAAG 240  
QY 81 ThrValGlnGluAlaGluAenIleAlaValThrSerGlyValAlaArgIysGluAspLeu 100  
Db 241 ACTGTGAGAGAGGTGGAGAACTTGCACTCACTGTGAGTGTGTCACAGAGAGCCCTG 300  
QY 101 ArgProSerAlaProGlnGlnGluIysGluIysSerIysGluIysGluValAlaGlu 120  
Db 301 AAGCACTGTGCCCCCAAGAGAGGTGGGCTATCCAAAGAAAGAGAGTGGCAGAG 360  
QY 121 GlnAlaGlnSerGlyIysAsp 127  
Db 361 GAGACCAAGAGTGGGGAGAT 381

## RESULT 15

E36334 727 bp DNA linear PAT 18-JUN-2001  
LOCUS Analytical matter based on synuclein and novel synuclein protein.  
DEFINITION E36334  
ACCESSION E36334.1 GI:13022627  
VERSION  
KEYWORDS JP 1999239488-A/2.  
SOURCE Mus musculus (house mouse)

REFERENCE  
AUTHORS Andrew, S.M., Valdimitra, R.B. and Arun, M.D.  
TITLE Analytical matter based on synuclein and novel synuclein protein  
JOURNAL Patent: JP 1999239488-A 2 07-SEP-1999;  
THE UNIVERSITY COURT OF THE UNIVERSITY OF ST ANDREWS, NYUROBA LTD  
OS Mus musculus (mouse)  
PN JP 1999239488-A/2  
PD 07-SEP-1999  
PF 21-SEP-1998 JP 1998306283  
PR 19-SEP-1997 GB 9719879.0  
PI ANDREW SMITH MAKARION, VALDIMITRA RUVOVICH BUCHIMAN, PI ARUN

COMMENT  
MILWARD DAVIS  
PC C12N15/09, A01K67/027, C12Q1/68, G01N33/53, C12N15/00 CC  
FH Key  
FT CDS location/Qualifiers (69)..(440).

FEATURES  
source location/Qualifiers

CDS 1..727  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"

## ORIGIN

Alignment Scores:  
Pred. No.: 4.93e-34 Length: 727  
Score: 497.00 Matches: 108  
Percent Similarity: 89.76% Conservative: 6

Best Local Similarity: 85.04% Mismatches: 9  
Query Match: 81.48% Indels: 4  
DB: 6 Gaps: 2

US-09-017-715A-2 (1-127) x E36334 (1-727)

```
OY      1 MetAspValPheLysLysGlyPheSerIleAlaLysLysGlyValValGlyAlaValGlu 20
      69 ATGAGCTCTTCAAGAAAAGGCTTCTCCATTGCCAAGAAAGTGTGTGTGCTGTAGAA 128
OY      21 LysThrLysGlnGlyValThrGluAlaAlaGluLysThrLysGluGlyValMetTyVal 40
      129 AAGACCAAGCAGGAGGAGTAAACGAGCGACCTGAGAAGACCAAGAGGGGTTATGTATGTG 188
OY      41 GlyAlaLysThrLysGluAsnValValGlnSerValThrSerValAlaGluLysThrLys 60
      189 GGCACCAAAACCAAGAGAGACGTGTACAAAGTGTCACTCAGTGGCTGAGAAAGACCAAG 248
OY      61 GluGlnAlaAsnAlaValSerLysAlaValValSerSerValAsnThrValAlaThrLys 80
      249 GAGCAGGCGCAATGCCGTGAGTGAAGCTGTGTCAAGCGCTCAACACAGTGGCCAAAG 308
OY      81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValArgLysGluAspLeu 100
      309 ACCGTGAGAGAGGCTGAAAAACATCGTGTCAACCAACCGGGGTGTGCGCAAGAGAGACTTG 368
OY      101 ArgProSerAlaProGlnGlnGlnGluGluAlaSerLysGluLysGluGluValAlaGlu 120
      369 GAACCCCTGCAACAGACCAAGAA-----GCCAAGAGCAAGAGGAG---AATGAA 416
DB      121 GluAlaGlnSerGlyGlyAsp 127
      417 GAGGCCAAGAGTGTGAGAGAGAC 437
```

Search completed: May 4, 2005, 11:53:21  
Job time : 7221.91 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 4, 2005, 10:17:21 ; Search time 65.0738 Seconds  
(without alignments)  
748,404 Million cell updates/sec

Title: US-09-017-715A-2\_COPY\_120\_127  
Perfect score: 41  
Sequence: 1 EEAQSGSD 8

Scoring table:  
BLOSUM62  
Xgapop 10.0, Ygapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 5642217 seqs, 3043843248 residues  
Total number of hits satisfying chosen parameters: 11284434

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cg21/USPTO.spool.h/US09017715/runat\_04052005\_100747\_25719/app\_query.fasta\_1.661  
-DB=Published Applications NA -QFMT=fastab -SUFFIX=emb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blone62  
-TRANS=human40.cct -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09017715 @CGN\_1\_1\_684 @runat\_04052005\_100747\_25719  
-NCPUPU=6 -ICPU=3 -NO MMAP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA: \*  
1: /cg21\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq: \*  
2: /cg21\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cg21\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cg21\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq: \*  
5: /cg21\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cg21\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cg21\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cg21\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq: \*  
9: /cg21\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq: \*  
10: /cg21\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq: \*  
11: /cg21\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq: \*  
12: /cg21\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq: \*  
13: /cg21\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq: \*  
14: /cg21\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq: \*  
15: /cg21\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq: \*  
16: /cg21\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq: \*  
17: /cg21\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq: \*  
18: /cg21\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq: \*  
19: /cg21\_6/ptodata/1/pubpna/US10H\_NEW\_PUB.seq: \*  
20: /cg21\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq: \*  
21: /cg21\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq: \*  
22: /cg21\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	384	19 US-10-826-157-5	Sequence 5, Appl1
2	41	100.0	478	9 US-09-925-298-172	Sequence 172, App
3	41	100.0	478	14 US-10-102-806-172	Sequence 172, App
4	41	100.0	478	10 US-09-918-995-2705	Sequence 2705, Ap
5	41	100.0	550	9 US-09-954-531-613	Sequence 613, App
6	41	100.0	550	17 US-10-453-478-1	Sequence 1, Appl1
7	41	100.0	550	19 US-10-843-641A-1680	Sequence 1680, Ap
8	41	100.0	720	14 US-10-097-340-297	Sequence 297, App
9	41	100.0	720	17 US-10-282-174-469	Sequence 469, App
10	41	100.0	720	19 US-10-600-009-469	Sequence 469, App
11	41	100.0	796	9 US-09-925-298-171	Sequence 171, App
12	41	100.0	796	14 US-10-102-806-171	Sequence 171, App
13	38	92.7	1125	17 US-10-369-493-33286	Sequence 33286, A
14	37	90.2	442	18 US-10-425-115-90466	Sequence 90466, A
15	37	90.2	2185	18 US-10-437-963-8217	Sequence 8217, Ap
16	37	90.2	106236	19 US-10-741-660-17759	Sequence 17759, A
17	36	87.8	357	18 US-10-437-963-43087	Sequence 43087, A
18	36	87.8	406	17 US-10-424-599-15323	Sequence 15323, A
19	36	87.8	1174	13 US-10-027-632-123222	Sequence 123222, A
20	36	87.8	1174	13 US-10-027-632-123222	Sequence 123223, A
21	36	87.8	1174	17 US-10-027-632-123222	Sequence 123223, A
22	36	87.8	1174	17 US-10-027-632-123223	Sequence 123223, A
23	36	87.8	2116	17 US-10-104-047-10939	Sequence 10939, Ap
24	36	87.8	3016	18 US-10-602-494-48	Sequence 48, Appl
25	36	87.8	4985	13 US-10-071-223-1	Sequence 1, Appl1
26	36	87.8	4985	19 US-10-896-552-1	Sequence 1897, Ap
27	36	87.8	5372	18 US-10-723-860-8187	Sequence 8187, Ap
28	36	87.8	5515	9 US-09-751-1008-98	Sequence 98, Appl
29	36	87.8	225666	18 US-10-470-565-1	Sequence 1, Appl1
30	35	85.4	291	18 US-10-425-115-65435	Sequence 65435, A
31	35	85.4	397	18 US-10-767-701-27423	Sequence 27423, A
32	35	85.4	518	18 US-10-021-323-3503	Sequence 3503, Ap
33	35	85.4	547	16 US-10-029-386-6547	Sequence 6547, Ap
34	35	85.4	588	18 US-10-437-963-6679	Sequence 6679, Ap
35	35	85.4	775	17 US-10-425-114-6521	Sequence 6521, Ap
36	35	85.4	775	18 US-10-425-115-10074	Sequence 10074, A
37	35	85.4	884	18 US-10-425-115-119416	Sequence 119416, A
38	35	85.4	918	17 US-10-369-493-44914	Sequence 44914, A
39	35	85.4	1273	14 US-10-198-846-12657	Sequence 12657, A
40	35	85.4	1672	18 US-10-425-115-145170	Sequence 145170, A
41	35	85.4	1836	17 US-10-424-599-129949	Sequence 129949, A
42	35	85.4	1885	18 US-10-437-963-75512	Sequence 75512, A
43	35	85.4	1901	17 US-10-424-599-131648	Sequence 131648, A
44	35	85.4	1928	17 US-10-425-114-24113	Sequence 24113, A
45	35	85.4	2047	17 US-10-425-114-27013	Sequence 27013, A

ALIGNMENTS

RESULT 1  
US-10-826-157-5  
Sequence 5, Application US/10826157  
Publication No. US20050064548A1  
GENERAL INFORMATION:  
APPLICANT: Lindquist, Susan L.  
TITLE OF INVENTION: YEAST ECTOPOICALLY EXPRESSING ABNORMALLY  
FILE REFERENCE: 17481-003001  
CURRENT APPLICATION NUMBER: US/10/826,157  
CURRENT FILING DATE: 2004-04-16  
PRIOR APPLICATION NUMBER: US 60/472,317  
PRIOR FILING DATE: 2003-05-20  
PRIOR APPLICATION NUMBER: US 60/463,284  
PRIOR FILING DATE: 2003-04-16  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 384

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-826-157-5

Alignment Scores:
Pred. No.: 20.2      Length: 384
Score: 41.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 19      Gaps: 0

US-09-017-715A-2_COPY_120_127 (1-8) x US-10-826-157-5 (1-384)
QY 1 GIUGUAlaGInSerGIyGIyAsp 8
Db 358 GAGGAGGCCAGAGTGGGGAGAC 381

RESULT 2
US-09-925-298-172
; Sequence 172, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 172
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-298-172

Alignment Scores:
Pred. No.: 24.4      Length: 478
Score: 41.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 9      Gaps: 0

US-09-017-715A-2_COPY_120_127 (1-8) x US-09-925-298-172 (1-478)
QY 1 GIUGUAlaGInSerGIyGIyAsp 8
Db 149 GAGGAGGCCAGAGTGGGGAGAC 172

RESULT 3
US-10-102-806-172
; Sequence 172, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 172
```

```
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-806-172

Alignment Scores:
Pred. No.: 24.4      Length: 478
Score: 41.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 14      Gaps: 0

US-09-017-715A-2_COPY_120_127 (1-8) x US-10-102-806-172 (1-478)
QY 1 GIUGUAlaGInSerGIyGIyAsp 8
Db 149 GAGGAGGCCAGAGTGGGGAGAC 172

RESULT 4
US-09-918-995-2705
; Sequence 2705, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2705
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(479)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-2705

Alignment Scores:
Pred. No.: 24.4      Length: 479
Score: 41.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 10      Gaps: 0

US-09-017-715A-2_COPY_120_127 (1-8) x US-09-918-995-2705 (1-479)
QY 1 GIUGUAlaGInSerGIyGIyAsp 8
Db 371 GAGGAGGCCAGAGTGGGGAGAC 394

RESULT 5
US-09-954-531-613
; Sequence 613, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cance
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
```



PRIOR APPLICATION NUMBER: US/60/234,034  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234,509  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: US/60/234,567  
PRIOR FILING DATE: 2000-09-22  
NUMBER OF SEQ ID NOS: 1392  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO: 613  
LENGTH: 550  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-954-531-613

Alignment Scores:  
Pred. No.: 27.5 Length: 550  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-017-715a-2\_COPY\_120\_127 (1-8) x US-09-954-531-613 (1-550)

Qy 1 GluGluAlaGlnSerGlyGlyAsp 8  
Db 369 GAGGAGGCCGAGCTGGGGGAGAC 392

RESULT 6  
US-10-453-478-1  
Sequence 1, Application US/10453478  
Publication No. US20030208043A1

GENERAL INFORMATION:  
APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,  
Jian Ni and Jing-Shan Hu  
TITLE OF INVENTION: Human Genes, Sequences and  
Expression Products  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/453,478  
FILING DATE: 04-Jun-2003  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/705,771  
FILING DATE: August 30, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: MULHINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-346 (PFI96)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 973-994-1700  
TELEFAX: 973-994-1744  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 550 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-453-478-1

Alignment Scores:  
Pred. No.: 27.5 Length: 550  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0

US-09-017-715a-2\_COPY\_120\_127 (1-8) x US-10-453-478-1 (1-550)

Qy 1 GluGluAlaGlnSerGlyGlyAsp 8  
Db 369 GAGGAGGCCGAGCTGGGGGAGAC 392

RESULT 7  
US-10-843-641A-1680  
Sequence 1680, Application US/10843641A  
Publication No. US20050064454A1

GENERAL INFORMATION:  
APPLICANT: Avalon Pharmaceuticals, Inc.  
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
Signature Gene Sets  
FILE REFERENCE: 689290-189  
CURRENT APPLICATION NUMBER: US/10/843,641A  
CURRENT FILING DATE: 2004-05-12  
PRIOR APPLICATION NUMBER: US/09/873,367  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: US/09/954,531  
PRIOR FILING DATE: 2001-09-18  
PRIOR APPLICATION NUMBER: US/09/954,456  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: US/09/962,436  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: US/09/962,832  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: US/09/964,824  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: US/09/967,768  
PRIOR FILING DATE: 2001-09-28  
PRIOR APPLICATION NUMBER: US/09/968,007  
PRIOR FILING DATE: 2001-10-02  
PRIOR APPLICATION NUMBER: US/09/969,347  
PRIOR FILING DATE: 2001-10-02  
PRIOR APPLICATION NUMBER: US/09/969,708  
PRIOR FILING DATE: 2001-10-03  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 8447  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO: 1680  
LENGTH: 550  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-843-641A-1680

Alignment Scores:  
Pred. No.: 27.5 Length: 550  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 19 Gaps: 0

US-09-017-715a-2\_COPY\_120\_127 (1-8) x US-10-843-641A-1680 (1-550)

Qy 1 GluGluAlaGlnSerGlyGlyAsp 8  
Db 369 GAGGAGGCCGAGCTGGGGGAGAC 392

RESULT 8  
US-10-097-340-297  
Sequence 297, Application US/10097340

```
/ Publication No. US20030087250A1
/ GENERAL INFORMATION:
/ APPLICANT: John MONAHAN
/ APPLICANT: Manjula GANNANARAPU
/ APPLICANT: Sebastian HOERSCH
/ APPLICANT: Shubhangi KAMATKAR
/ APPLICANT: Steve G. KOVATS
/ APPLICANT: Rachel E. MEYERS
/ APPLICANT: Michael MORRISSEY
/ APPLICANT: Peter OLANDT
/ APPLICANT: Ami SEN
/ APPLICANT: Peter VEIBY
/ APPLICANT: Gordon B. MILLS
/ APPLICANT: Robert C. BAST, Jr.
/ APPLICANT: Karen LU
/ APPLICANT: Rosemarie SCHMANDT
/ APPLICANT: Xumei ZHAO
/ APPLICANT: Karen GIATT
/ TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
/ FILE REFERENCE: MRI-030
/ TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
/ CURRENT APPLICATION NUMBER: US/10/097,340
/ CURRENT FILING DATE: 2002-03-14
/ PRIOR APPLICATION NUMBER: 60/276,025
/ PRIOR FILING DATE: 2001-03-14
/ PRIOR APPLICATION NUMBER: 60/325,149
/ PRIOR FILING DATE: 2001-09-26
/ PRIOR APPLICATION NUMBER: 60/276,026
/ PRIOR FILING DATE: 2001-03-14
/ PRIOR APPLICATION NUMBER: 60/324,967
/ PRIOR FILING DATE: 2001/09/26
/ PRIOR APPLICATION NUMBER: 60/311,732
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: 60/325,102
/ PRIOR FILING DATE: 2001-09-26
/ PRIOR APPLICATION NUMBER: 60/323,580
/ PRIOR FILING DATE: 2001-09-19
/ NUMBER OF SEQ ID NOS: 363
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 297
/ LENGTH: 720
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-097-340-297

Alignment Scores:
Pred. No.: 34.8      Length: 720
Score: 41.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 14      Gaps: 0

US-09-017-715a-2_copy_120_127 (1-8) x US-10-097-340-297 (1-720)
Qy 1 GIUGUAAGInSerGIyGIyASP 8
Db 406 GAGGAGCCCCAGAGTGGGGGAGAC 429

RESULT 9
US-10-282-174-469
/ Sequence 469, Application US/10282174
/ Publication No. US20030224380A1
/ GENERAL INFORMATION:
/ APPLICANT: Becker, Kenneth David
/ APPLICANT: Velicelebi, Gonul
/ APPLICANT: Eliot, Kathryn J.
/ APPLICANT: Wang, Xin
/ APPLICANT: Tanzi, Rudolph E.
/ APPLICANT: Bertam, Lars
/ APPLICANT: Saunders, Aleister J.
/ APPLICANT: Mullin, Kristina M.
/ APPLICANT: Sampson, Andrew Johnson
```

```
/ APPLICANT: Blacker, Deborah Lynne
/ TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
/ TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
/ TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
/ FILE REFERENCE: 37481-3308
/ CURRENT APPLICATION NUMBER: US/10/282,174
/ CURRENT FILING DATE: 2002-10-25
/ PRIOR APPLICATION NUMBER: US 60/339,525
/ PRIOR FILING DATE: 2001-10-25
/ PRIOR APPLICATION NUMBER: US 60/338,010
/ PRIOR FILING DATE: 2001-11-08
/ PRIOR APPLICATION NUMBER: US 60/336,929
/ PRIOR FILING DATE: 2001-11-08
/ PRIOR APPLICATION NUMBER: US 60/338,363
/ PRIOR FILING DATE: 2001-11-09
/ PRIOR APPLICATION NUMBER: US 60/337,052
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: US 60/368,919
/ PRIOR FILING DATE: 2002-03-28
/ NUMBER OF SEQ ID NOS: 564
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 469
/ LENGTH: 720
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 30,57,85,243,250,377,512,531,555,561,672
/ OTHER INFORMATION: N is any
/ US-10-282-174-469

Alignment Scores:
Pred. No.: 34.8      Length: 720
Score: 41.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 17      Gaps: 0

US-09-017-715a-2_copy_120_127 (1-8) x US-10-282-174-469 (1-720)
Qy 1 GIUGUAAGInSerGIyGIyASP 8
Db 406 GAGGAGCCCCAGAGTGGGGGAGAC 429

RESULT 10
US-10-600-009-469
/ Sequence 469, Application US/10600009
/ Publication No. US2005009031A1
/ GENERAL INFORMATION:
/ APPLICANT: Becker, Kenneth David
/ APPLICANT: Velicelebi, Gonul
/ APPLICANT: Eliot, Kathryn J.
/ APPLICANT: Wang, Xin
/ APPLICANT: Tanzi, Rudolph E.
/ APPLICANT: Bertam, Lars
/ APPLICANT: Saunders, Aleister J.
/ APPLICANT: Mullin, Kristina M.
/ APPLICANT: Sampson, Andrew Johnson
/ APPLICANT: Blacker, Deborah Lynne
/ TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
/ TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
/ TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
/ FILE REFERENCE: 37481-3308B
/ CURRENT APPLICATION NUMBER: US/10/600,009
/ CURRENT FILING DATE: 2003-06-18
/ PRIOR APPLICATION NUMBER: US 60/339,525
/ PRIOR FILING DATE: 2001-10-25
/ PRIOR APPLICATION NUMBER: US 60/338,010
/ PRIOR FILING DATE: 2001-11-08
/ PRIOR APPLICATION NUMBER: US 60/336,929
/ PRIOR FILING DATE: 2001-11-08
/ PRIOR APPLICATION NUMBER: US 60/338,363
```

```

; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/337,052
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/368,919
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 10/282,174
; PRIOR FILING DATE: 2002-10-25
; NUMBER OF SEQ ID NOS: 564
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 469
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 30,57,85,243,250,377,512,531,555,561,672
; OTHER INFORMATION: N is any
; US-10-600-009-469

```

```

Alignment Scores:
Pred. No.: 34.8      Length: 720
Score: 41.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 19      Gaps: 0

```

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-10-600-009-469 (1-720)

QY 1 GluGluaIaGInserGIyGIyAaP 8

Db 406 GAGGAGGCCAGAGTGGGGAGAC 429

```

RESULT 11
US-09-925-298-171
; Sequence 171, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 171
; LENGTH: 796
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-298-171

```

```

Alignment Scores:
Pred. No.: 37.9      Length: 796
Score: 41.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 9      Gaps: 0

```

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-09-925-298-171 (1-796)

QY 1 GluGluaIaGInserGIyGIyAaP 8

Db 466 GAGGAGGCCAGAGTGGGGAGAC 489

```

RESULT 12
US-10-102-806-171
; Sequence 171, Application US/10102806
; Publication No. US20030054421A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 171
; LENGTH: 796
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-102-806-171

```

```

Alignment Scores:
Pred. No.: 37.9      Length: 796
Score: 41.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 14      Gaps: 0

```

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-10-102-806-171 (1-796)

QY 1 GluGluaIaGInserGIyGIyAaP 8

Db 466 GAGGAGGCCAGAGTGGGGAGAC 489

```

RESULT 13
US-10-369-493-33286/c
; Sequence 33286, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 33286
; LENGTH: 1125
; TYPE: DNA
; ORGANISM: Desulfotobacterium hafnienae
; US-10-369-493-33286

```

```

Alignment Scores:
Pred. No.: 201      Length: 1125
Score: 38.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 87.50%      Mismatches: 0
Query Match: 92.68%      Indels: 0
DB: 17      Gaps: 0

```

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-10-369-493-33286 (1-1125)

QY 1 GluGluaIaGInserGIyGIyAaP 8

Db 609 GAGGAGGCCAGAGTGGGGAGAC 586

RESULT 14

US-10-425-115-90466/c  
; Sequence 90466, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 90466  
; LENGTH: 442  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(442)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_182507C.1  
US-10-425-115-90466

Alignment Scores:  
Pred. No.: 142 Length: 442  
Score: 37.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 87.50% Mismatches: 0  
Query Match: 90.24% Indels: 0  
DB: 18 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-10-425-115-90466 (1-442)

Qy 1 GUGUUAAGInserGlyGlyASP 8  
DB 55 GAGGAGCGAGGAGCGAGGAGC 32

RESULT 15  
US-10-437-963-8217  
; Sequence 8217, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 8217  
; LENGTH: 2185  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_14742C.1  
US-10-437-963-8217

Alignment Scores:  
Pred. No.: 564 Length: 2185  
Score: 37.00 Matches: 7  
Percent Similarity: 87.50% Conservative: 0  
Best Local Similarity: 87.50% Mismatches: 1  
Query Match: 90.24% Indels: 0

DB: 18 Gaps: 0  
US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-10-437-963-8217 (1-2185)  
Qy 1 GUGUUAAGInserGlyGlyASP 8  
DB 1208 GAGGAGCTCAGGAGGAGGTGAC 1231

Search completed: May 4, 2005, 16:39:46  
Job time : 68.7405 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 4, 2005, 09:07:55 ; Search time 18.5235 Seconds  
(without alignments)  
706.682 Million cell updates/sec

Title: US-09-017-715a-2\_COPY\_120\_127

Perfect score: 41  
Sequence: 1 EEAQSGSD 8

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool\_h/US09017715/runat\_04052005\_100745\_25632/app\_query.fasta\_1.661  
-DB=Issued\_Patents\_NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPTC=0  
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=p2n -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US09017715\_@CGN\_1\_1\_116\_@runat\_04052005\_100745\_25632 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	41	100.0	550	3	US-08-705-771-1
2	41	100.0	550	4	US-09-417-540-1
3	41	100.0	702	4	US-09-949-016-1915
4	41	100.0	720	4	US-09-949-016-442
5	36	87.8	1902	4	US-09-949-016-5550
6	36	87.8	4985	4	US-09-473-717-1
7	36	87.8	4985	4	US-09-949-016-152
8	36	87.8	5515	3	US-09-398-193-98
9	36	87.8	112114	4	US-09-949-016-17292
10	36	87.8	126200	4	US-09-949-016-11824
11	36	87.8	126200	4	US-09-949-016-13193
12	36	87.8	154605	4	US-09-949-016-11894

C 13	35	85.4	83708	4	US-09-949-016-17207	Sequence 17207, A
C 14	34	82.9	601	4	US-09-949-016-49739	Sequence 49739, A
C 15	34	82.9	628	3	US-09-328-111-367	Sequence 367, App
C 16	34	82.9	1653	4	US-09-830-111E-1	Sequence 1, Appl1
C 17	34	82.9	29717	4	US-09-949-016-16284	Sequence 16284, A
C 18	34	82.9	60376	4	US-09-949-016-12423	Sequence 12423, A
C 19	33	80.5	204	4	US-09-621-976-18715	Sequence 18715, A
C 20	33	80.5	210	4	US-09-621-976-18727	Sequence 18727, A
C 21	33	80.5	219	4	US-09-621-976-18782	Sequence 18782, A
C 22	33	80.5	229	4	US-09-621-976-18788	Sequence 18788, A
C 23	33	80.5	361	4	US-09-621-976-18728	Sequence 18728, A
C 24	33	80.5	404	4	US-09-621-976-18769	Sequence 18769, A
C 25	33	80.5	431	4	US-09-513-999C-13044	Sequence 13044, A
C 26	33	80.5	460	4	US-09-621-976-3657	Sequence 3657, App
C 27	33	80.5	480	4	US-09-270-767-6338	Sequence 6338, App
C 28	33	80.5	480	4	US-09-270-767-21620	Sequence 21620, A
C 29	33	80.5	487	4	US-09-621-976-18762	Sequence 18762, A
C 30	33	80.5	495	4	US-09-621-976-3659	Sequence 3659, App
C 31	33	80.5	601	4	US-09-949-016-26549	Sequence 26549, A
C 32	33	80.5	601	4	US-09-949-016-26550	Sequence 26550, A
C 33	33	80.5	601	4	US-09-949-016-27764	Sequence 27764, A
C 34	33	80.5	601	4	US-09-949-016-65608	Sequence 65608, A
C 35	33	80.5	601	4	US-09-949-016-65609	Sequence 65609, A
C 36	33	80.5	601	4	US-09-949-016-164928	Sequence 164928, A
C 37	33	80.5	601	4	US-09-949-016-202311	Sequence 202311, A
C 38	33	80.5	670	3	US-09-404-879A-156	Sequence 156, App
C 39	33	80.5	670	4	US-09-338-933-156	Sequence 156, App
C 40	33	80.5	670	4	US-09-215-681-156	Sequence 156, App
C 41	33	80.5	670	4	US-09-216-003A-156	Sequence 156, App
C 42	33	80.5	670	4	US-09-667-857-156	Sequence 156, App
C 43	33	80.5	978	4	US-09-248-796A-5309	Sequence 5309, App
C 44	33	80.5	1005	4	US-09-270-767-26942	Sequence 26942, A
C 45	33	80.5	1208	4	US-09-270-767-25543	Sequence 25543, A

## ALIGNMENTS

RESULT 1  
US-08-705-771-1  
Sequence 1, Application US/08705771  
Patent No. 6054289  
GENERAL INFORMATION:  
APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,  
TITLE OF INVENTION: Human Genes, Sequences and  
TITLE OF INVENTION: Expression Products  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/705, 771  
FILING DATE: August 30, 1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33, 073  
REFERENCE/DOCKET NUMBER: 325800-346 (P1196)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 973-994-1744  
TELEFAX: 973-994-1744  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 550 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-705-771-1

Alignment Scores:  
Pred. No.: 9.29 Length: 550  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-08-705-771-1 (1-550)

Qy 1 GUGUUAAGInserG1yG1yasp 8  
Db 369 GAGGAGCCCGAGGTGGGGAGAC 392

RESULT 2  
US-09-417-540-1  
Sequence 1, Application US/09417540  
Patent No. 6639052  
GENERAL INFORMATION:  
APPLICANT: Moore, Reiner Gentz, Hongjin Ji,  
Jian Ni and Jing-Shan Hu  
TITLE OF INVENTION: Human Genes, Sequences and  
Expression Products  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CECCHI, STEWART & OLSTEIN,  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/417,540  
FILING DATE: 14-OCT-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/705,771  
FILING DATE: August 30, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-346 (PFI96)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 973-994-1700  
TELEFAX: 973-994-1744  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 550 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-417-540-1

Alignment Scores:  
Pred. No.: 9.29 Length: 550  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-09-417-540-1 (1-550)

Qy 1 GUGUUAAGInserG1yG1yasp 8  
Db 369 GAGGAGCCCGAGGTGGGGAGAC 392

RESULT 3  
US-09-949-016-1915  
Sequence 1915, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1915  
LENGTH: 702  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-1915

Alignment Scores:  
Pred. No.: 12.1 Length: 702  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-09-949-016-1915 (1-702)

Qy 1 GUGUUAAGInserG1yG1yasp 8  
Db 405 GAGGAGCCCGAGGTGGGGAGAC 428

RESULT 4  
US-09-949-016-442  
Sequence 442, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 442  
LENGTH: 720  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-442

## Alignment Scores:

Pred. No.: 12, 4 Length: 720  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-09-949-016-442 (1-720)

QY 1 GluGluAlaGlnSerGlyIasp 8

DB 406 GAGGAGGCCAGAGTGGGGGAGAC 429

## RESULT 5

US-09-949-016-5550/C  
Sequence 5550, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949, 016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5550  
LENGTH: 1902  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-5550

## Alignment Scores:

Pred. No.: 366 Length: 1902  
Score: 36.00 Matches: 7  
Percent Similarity: 87.50% Conservative: 0  
Best Local Similarity: 87.50% Mismatches: 1  
Query Match: 87.80% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-09-949-016-5550 (1-1902)

QY 1 GluGluAlaGlnSerGlyIasp 8

DB 477 GAGGAGGCCAGAGTGGGGGAGAC 454

## RESULT 6

US-09-473-717-1/C  
Sequence 1, Application US/09473717  
Patent No. 6372475  
GENERAL INFORMATION:  
APPLICANT: Storm, Daniel R.  
APPLICANT: Hacker, Beth  
APPLICANT: Tomlinson, James E.  
APPLICANT: COR Therapeutics, Inc.  
TITLE OF INVENTION: Cloning and characterization of a human adenylate  
cyclase  
FILE REFERENCE: 44481-5029-01-US  
CURRENT APPLICATION NUMBER: US/09/473, 717  
CURRENT FILING DATE: 1999-12-29  
PRIOR APPLICATION NUMBER: PCT/US98/13541  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/098,559  
PRIOR FILING DATE: 1997-07-01  
PRIOR APPLICATION NUMBER: 08/886,440

PRIOR FILING DATE: 1997-07-01

NUMBER OF SEQ ID NOS: 3

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 4985

TYPE: DNA

ORGANISM: human type IX adenylate cyclase

FEATURE:

NAME/KEY: CDS

LOCATION: (17)..(3898)

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-09-473-717-1 (1-4985)

QY 1 GluGluAlaGlnSerGlyIasp 8

DB 475 GAGGAGGCCAGAGTGGGGGAGAC 452

## RESULT 7

US-09-949-016-152/C  
Sequence 152, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949, 016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 152  
LENGTH: 4985  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-152

## Alignment Scores:

Pred. No.: 1,046+03 Length: 4985  
Score: 36.00 Matches: 7  
Percent Similarity: 87.50% Conservative: 0  
Best Local Similarity: 87.50% Mismatches: 1  
Query Match: 87.80% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-09-949-016-152 (1-4985)

QY 1 GluGluAlaGlnSerGlyIasp 8

DB 475 GAGGAGGCCAGAGTGGGGGAGAC 452

## RESULT 8

US-09-398-193-98/C  
Sequence 98, Application US/09398193  
Patent No. 6197581  
GENERAL INFORMATION:  
APPLICANT: Medical Research Council  
TITLE OF INVENTION: Adenylate cyclase and uses therefor  
FILE REFERENCE: P24360-

```

; CURRENT APPLICATION NUMBER: US/09/398,193
; CURRENT FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 5515
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (539)..(4600)
; US-09-398-193-98

Alignment Scores:
Pred. No.: 1,16e+03 Length: 5515
Score: 36.00 Matches: 7
Percent Similarity: 87.50% Conservative: 0
Best Local Similarity: 87.50% Mismatches: 1
Query Match: 87.80% Indels: 0
DB: 3 Gaps: 0

US-09-017-715a-2_copy_120_127 (1-8) x US-09-398-193-98 (1-5515)

QY 1 GUGUAGAGlnSerGlyGlyASP 8
DB 997 GAGGAGCACACAGCGCGGGCGGAC 974

RESULT 9
US-09-949-016-17292/c
; Sequence 17292, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17292
; LENGTH: 112114
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-17292

Alignment Scores:
Pred. No.: 2.98e+04 Length: 112114
Score: 36.00 Matches: 7
Percent Similarity: 87.50% Conservative: 0
Best Local Similarity: 87.50% Mismatches: 1
Query Match: 87.80% Indels: 0
DB: 4 Gaps: 0

US-09-017-715a-2_copy_120_127 (1-8) x US-09-949-016-17292 (1-112114)

QY 1 GUGUAGAGlnSerGlyGlyASP 8
DB 2477 GAGGAGCACACAGCGCGGGCGGAC 2454

RESULT 10
US-09-949-016-11824/c
; Sequence 11824, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11824
; LENGTH: 126200
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(126200)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-11824

Alignment Scores:
Pred. No.: 3.39e+04 Length: 126200
Score: 36.00 Matches: 7
Percent Similarity: 87.50% Conservative: 0
Best Local Similarity: 87.50% Mismatches: 1
Query Match: 87.80% Indels: 0
DB: 4 Gaps: 0

US-09-017-715a-2_copy_120_127 (1-8) x US-09-949-016-11824 (1-126200)

QY 1 GUGUAGAGlnSerGlyGlyASP 8
DB 110246 GAGGAGCACACAGTGCGGGCGGAT 110223

RESULT 11
US-09-949-016-13193/c
; Sequence 13193, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13193
; LENGTH: 126200
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(126200)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-13193

Alignment Scores:
Pred. No.: 3.39e+04 Length: 126200
Score: 36.00 Matches: 7
Percent Similarity: 87.50% Conservative: 0
Best Local Similarity: 87.50% Mismatches: 1
Query Match: 87.80% Indels: 0
DB: 4 Gaps: 0
```



US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-09-949-016-13193 (1-126200)

Qy 1 GluGluaLagInserGlyGlyAap 8  
Db 110246 GAGGAGGACACAGCTGGGGGGAT 110223

RESULT 12

US-09-949-016-11894/C  
; Sequence 11894, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11894  
; LENGTH: 154605  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-11894

Alignment Scores:

Pred. No.:	4.21e+04	Length:	154605
Score:	36.00	Matches:	7
Percent Similarity:	87.50%	Conservative:	0
Best Local Similarity:	87.50%	Mismatches:	1
Query Match:	87.80%	Indels:	0
DB:	4	Gaps:	0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-09-949-016-11894 (1-154605)

Qy 1 GluGluaLagInserGlyGlyAap 8  
Db 2475 GAGGAGGACACAGCTGGGGGGAT 2452

RESULT 13

US-09-949-016-17207/C  
; Sequence 17207, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17207  
; LENGTH: 83708  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17207

Alignment Scores:

Pred. No.:	3.46e+04	Length:	83708
Score:	35.00	Matches:	7

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	85.37%	Indels:	0
DB:	4	Gaps:	0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-09-949-016-17207 (1-83708)

Qy 1 GluGluaLagInserGlyGly 7  
Db 2492 GAAGAGGACACAGCTAGGGGCG 2472

RESULT 14

US-09-949-016-49739/C  
; Sequence 49739, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 49739  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-49739

Alignment Scores:

Pred. No.:	268	Length:	601
Score:	34.00	Matches:	7
Percent Similarity:	87.50%	Conservative:	0
Best Local Similarity:	87.50%	Mismatches:	1
Query Match:	82.93%	Indels:	0
DB:	4	Gaps:	0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x US-09-949-016-49739 (1-601)

Qy 1 GluGluaLagInserGlyGlyAap 8  
Db 56 GAGGTGCACACAGCTGGGGGAT 33

RESULT 15

US-09-328-111-367  
; Sequence 367, Application US/09328111  
; Patent No. 626233  
; GENERAL INFORMATION:  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Steinmann, Kathleen E.  
; APPLICANT: Ascle, Jon H.  
; APPLICANT: Burgess, Christopher C.  
; APPLICANT: Bushnell, Steven E.  
; APPLICANT: Carroll III, Eddie  
; APPLICANT: Catino, Theodore J.  
; APPLICANT: Dertl, Adnan  
; APPLICANT: Ford, Donna M.  
; APPLICANT: Lewis, Marcia E.  
; APPLICANT: Monahan, John E.  
; APPLICANT: Schlegel, Robert  
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
; FILE REFERENCE: CCD-257 (US)  
; CURRENT APPLICATION NUMBER: US/09/328,111  
; CURRENT FILING DATE: 1999-06-08  
; EARLIER APPLICATION NUMBER: US 60/088,801



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 4, 2005, 10:17:21 ; Search time 1033.05 Seconds  
(without alignments)  
748,404 Million cell updates/sec

Title: US-09-017-715A-2

Perfect score: 610  
Sequence: 1 MDVFKKGFSLAKKGVGAVE.....EGASKEKEVEAEASQSGD 127

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5642217 seqs, 3043843248 residues

Total number of hits satisfying chosen parameters: 11284434

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+g2n.model -DEV=xjh  
-Q=/cgn2\_1/USPRO.spool.h/US0901715/runat\_04052005\_100747\_25719/app\_query.fasta.1.661  
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US0901715@cgn.1.1.684@runat\_04052005\_100747\_25719  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DISPATCH=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
Published Applications NA: \*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq: \*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq: \*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq: \*  
14: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq: \*  
15: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq: \*  
16: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq: \*  
17: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq: \*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq: \*  
19: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq: \*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq: \*  
21: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq: \*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	610	100.0	550	9 US-09-954-531-613	Sequence 613, App
2	610	100.0	550	19 US-10-453-478-1	Sequence 1, Appl
3	610	100.0	550	19 US-10-843-641A-1680	Sequence 1680, Ap
4	602	98.7	796	4 US-09-925-286-171	Sequence 171, App
5	602	98.7	796	14 US-10-102-806-171	Sequence 171, App
6	595	97.5	384	19 US-10-826-157-5	Sequence 5, Appl
7	595	97.5	720	14 US-10-097-430-297	Sequence 297, App
8	592	97.0	720	19 US-10-282-174-469	Sequence 469, App
9	592	97.0	720	19 US-10-600-009-469	Sequence 469, App
10	538	88.2	479	10 US-09-918-995-2705	Sequence 2705, Ap
11	468.5	76.8	786	14 US-10-267-849-1	Sequence 1, Appl
12	328	53.8	210	18 US-10-204-337A-5	Sequence 1, Appl
13	316	51.8	1018	17 US-10-152-319A-1710	Sequence 1710, Ap
14	316	51.8	1018	19 US-10-486-706-260	Sequence 260, App
15	312.5	51.2	437	18 US-10-737-262-2	Sequence 2, Appl
16	307.5	50.4	423	14 US-10-077-584-1	Sequence 1, Appl
17	307.5	50.4	423	18 US-10-204-337A-3	Sequence 3, Appl
18	307.5	50.4	423	19 US-10-826-157-1	Sequence 1, Appl
19	307.5	50.4	437	18 US-10-737-262-1	Sequence 1, Appl
20	307.5	50.4	1105	15 US-10-223-978-10	Sequence 10, Appl
21	307.5	50.4	1466	15 US-10-101-510-362	Sequence 362, App
22	307.5	50.4	1543	18 US-10-721-693-14	Sequence 14, Appl
23	307.5	50.4	1543	18 US-10-852-997-11	Sequence 14, Appl
24	304.5	49.9	755	17 US-10-112-944-11	Sequence 11, Appl
25	304	49.8	441	9 US-09-960-352-12619	Sequence 12619, A
26	304	49.8	453	9 US-09-960-352-5029	Sequence 5029, Ap
27	303.5	49.8	1096	10 US-09-921-406C-35	Sequence 35, Appl
28	303.5	49.8	1096	18 US-10-721-693-23	Sequence 23, Appl
29	303.5	49.8	1096	18 US-10-852-997-23	Sequence 23, Appl
30	302	49.5	405	19 US-10-826-157-3	Sequence 3, Appl
31	253	41.5	249	18 US-10-204-337A-4	Sequence 4, Appl
32	243	39.8	424	9 US-09-960-352-1978	Sequence 1978, Ap
33	233	38.2	473	10 US-09-918-995-1832	Sequence 1832, Ap
34	205	33.6	456	10 US-09-918-995-26977	Sequence 26977, A
35	202	33.1	466	9 US-09-960-352-3369	Sequence 3369, Ap
36	201.5	33.0	393	17 US-10-240-425-3388	Sequence 388, App
37	201.5	33.0	5666	17 US-10-282-174-72	Sequence 72, Appl
38	201.5	33.0	5666	19 US-10-600-009-72	Sequence 72, Appl
39	201.5	33.0	6012	17 US-10-282-174-483	Sequence 483, App
40	201.5	33.0	6012	19 US-10-600-009-483	Sequence 483, App
41	199	32.6	1332	16 US-10-029-386-25796	Sequence 25796, A
42	199	32.6	502	16 US-10-029-386-12096	Sequence 12096, A
43	197.5	32.4	5666	17 US-10-282-174-73	Sequence 73, Appl
44	197.5	32.4	5666	19 US-10-600-009-73	Sequence 73, Appl
45	195	32.0	521	16 US-10-029-386-9757	Sequence 9757, Ap

## ALIGNMENTS

RESULT 1  
US-09-954-531-613  
Sequence 613, Application US/09954531  
Patent No. US20020165180A1  
GENERAL INFORMATION:  
APPLICANT: Weaver, Zoe  
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc  
FILE REFERENCE: 689290-77  
CURRENT FILING DATE: 2002-05-02  
PRIOR APPLICATION NUMBER: US/09/954, 531  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US/60/233, 133  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234, 009  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234, 034  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234, 509  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: US/60/234, 567

PRIOR FILING DATE: 2000-09-22  
NUMBER OF SEQ ID NOS: 1392  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 613  
LENGTH: 550  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-954-531-613

Alignment Scores:  
Pred. No.: 2,15e-60 Length: 550  
Score: 610.00 Matches: 127  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-017-715a-2 (1-127) x US-09-954-531-613 (1-550)

QY 1 MetAspValPheLysGlyPheSerIleAlaLysGlyValValGluValGlu 20  
DB 12 ATGGATGTTTCAAGAGGGCTTCTCCATCGCCAGAGGGGCTGTGGTGGCGGGA 71

QY 21 LysThrLysGlnGlyValThrGluAlaGluLysThrLysGluGlyValMetCysVal 40  
DB 72 AAGACCAAGCAGGGGTGACGAGACGCTGAGAACCAAGAGGGGTCTATGTATGTG 131

QY 41 GlyAlaLysThrLysGluAsnValValGlnSerValThrSerValAlaGluLysThrLys 60  
DB 132 GGAGGCCAAGACCAAGAGATGTTGTACAGAGCGTACCTCGGTGCGCAAGAACCAAG 191

QY 61 GluGlnAlaAsnAlaValSerLysAlaValLysSerValAsnThrValAlaThrLys 80  
DB 192 GAGCAGGCCAAGCGCTGAGCAAGGCTGTGTGAGAGGCTCAACACTGTGGCCACCAAG 251

QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValArgLysGluAspLeu 100  
DB 252 ACCGTGAGAGAGCGGAGAACATCGCGTACCTCGGGGTGTGTGCGCAAGAGCACTTG 311

QY 101 ArgProSerAlaProGlnGlnGluGluLysGluLysGluGluValAlaGlu 120  
DB 312 AGGCCATCTGCCCCCAACAGAGGGGTGAGGCTCCAAAGAGAAAGAGAGTGGCAGAG 371

QY 121 GluAlaGlnSerGlyGlyAsp 127  
DB 372 GAGGCCCAAGAGTGGGGAGAGAC 392

RESULT 2  
US-10-453-478-1  
Sequence 1, Application US/10453478  
Publication No. US20030208043A1  
GENERAL INFORMATION:  
APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,  
Jian Ni and Jing-Shan Hu  
TITLE OF INVENTION: Human Genes, Sequences and  
Expression Products  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARRELLA, BYRNE, BAIN, GILFILLAN,  
CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/453,478  
FILING DATE: 04-Jun-2003

CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/705,771  
FILING DATE: August 30, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-346 (PFI96)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 973-994-1700  
TELEFAX: 973-994-1744  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 550 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-453-478-1

Alignment Scores:  
Pred. No.: 2,15e-60 Length: 550  
Score: 610.00 Matches: 127  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0

US-09-017-715a-2 (1-127) x US-10-453-478-1 (1-550)

QY 1 MetAspValPheLysGlyPheSerIleAlaLysGlyValValGluValGlu 20  
DB 12 ATGGATGTTTCAAGAGGGCTTCTCCATCGCCAGAGGGGCTGTGGTGGCGGGA 71

QY 21 LysThrLysGlnGlyValThrGluAlaGluLysThrLysGluGlyValMetCysVal 40  
DB 72 AAGACCAAGCAGGGGTGACGAGACGCTGAGAACCAAGAGGGGTCTATGTATGTG 131

QY 41 GlyAlaLysThrLysGluAsnValValGlnSerValThrSerValAlaGluLysThrLys 60  
DB 132 GGAGGCCAAGACCAAGAGATGTTGTACAGAGCGTACCTCGGTGCGCAAGAACCAAG 191

QY 61 GluGlnAlaAsnAlaValSerLysAlaValLysSerValAsnThrValAlaThrLys 80  
DB 192 GAGCAGGCCAAGCGCTGAGCAAGGCTGTGTGAGAGGCTCAACACTGTGGCCACCAAG 251

QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValArgLysGluAspLeu 100  
DB 252 ACCGTGAGAGAGCGGAGAACATCGCGTACCTCGGGGTGTGTGCGCAAGAGCACTTG 311

QY 101 ArgProSerAlaProGlnGlnGluGluLysGluLysGluGluValAlaGlu 120  
DB 312 AGGCCATCTGCCCCCAACAGAGGGGTGAGGCTCCAAAGAGAAAGAGAGTGGCAGAG 371

QY 121 GluAlaGlnSerGlyGlyAsp 127  
DB 372 GAGGCCCAAGAGTGGGGAGAGAC 392

RESULT 3  
US-10-843-641A-1680  
Sequence 1680, Application US/10843641A  
Publication No. US20050064454A1  
GENERAL INFORMATION:  
APPLICANT: Avalon Pharmaceuticals, Inc.  
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
File Reference: 689290-189  
FILE REFERENCE: Signature Gene Sets  
CURRENT APPLICATION NUMBER: US/10/843,641A  
CURRENT FILING DATE: 2004-05-12  
PRIOR APPLICATION NUMBER: US/09/873,367  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: US/09/954,531

```

; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1680
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-1680

Alignment Scores:
Pred. No.: 2,15e-60 Length: 550
Score: 610.00 Matches: 127
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19 Indels: 0
Gaps: 0

US-09-017-715A-2 (1-127) x US-10-843-641A-1680 (1-550)
QY 1 MetAspValPheIysLysGlyPheSerIleAlaIysLysGlyValAlaGlyAlaValAlaGlu 20
Db 12 ATGGATGTTTTCAGAAAGGGCTTTCATCCCAAGAGGGCGTGTGGTGGCGGTGAA 71
QY 21 LysThrLysGlnGlyValAlaThrGluAlaAlaGluLysThrLysGlnGlyValAlaMetLysVal 40
Db 72 AAGGCCAAGCAGGGGGGTGAGGAGAGCAGCTGAGAGAGCAAGAGAGGGGGTCATGTATGTG 131
QY 41 GlyAlaLysThrLysGluAsnValAlaGlnSerValThrSerValAlaGluLysThrLys 60
Db 132 GAGGCCAAGCAGGAGAAATGTTGTACAGAGCTGACCTCAGTGGCCGAGAGACCAAG 191
QY 61 GluGlnAlaAsnAlaValSerLysAlaValSerSerValAsnThrValAlaThrLys 80
Db 192 GAGCAGGCGCAACCGCGTGTGAGCAGGCTGAGCACTGAGCACTGTGGCCACCAAG 251
QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValAlaArgLysGluAspLeu 100
Db 252 ACCGTGAGAGAGCGCGAAGCATCGCGTCACTCCGGGGGTGGTCCCAAGAGAGACTTG 311
QY 101 ArgProSerAlaProGlnGlnGlnGlyGluAlaSerLysGluLysGlnGluValAlaGlu 120
Db 312 AGGCATCTGCCCCCAAGAGAGGGTGAAGCATCCAAAGAGAAAGAGAAAGTGGCAGAG 371
QY 121 GluAlaGlnSerGlyGlyAsp 127
Db 372 GAGGCCAAGTGGGGAGAC 392

RESULT 4
; Sequence 171, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298

```

```

; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 171
; LENGTH: 796
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-298-171

Alignment Scores:
Pred. No.: 2.87e-59 Length: 796
Score: 602.00 Matches: 125
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 98.43% Mismatches: 0
Query Match: 98.69% Indels: 0
Gaps: 0

US-09-017-715A-2 (1-127) x US-09-925-298-171 (1-796)
QY 1 MetAspValPheIysLysGlyPheSerIleAlaIysLysGlyValAlaGlyAlaValAlaGlu 20
Db 109 ATGGATGTTTTCAGAAAGGGCTTTCATCCCAAGAGGGCGTGTGGTGGCGGTGAA 168
QY 21 LysThrLysGlnGlyValAlaThrGluAlaAlaGluLysThrLysGlnGlyValAlaMetLysVal 40
Db 169 AAGGCCAAGCAGGGGGGTGAGGAGAGCAGCTGAGAGAGCAAGAGAGGGGGTCATGTATGTG 228
QY 41 GlyAlaLysThrLysGluAsnValAlaGlnSerValThrSerValAlaGluLysThrLys 60
Db 229 GAGGCCAAGCAGGAGAAATGTTGTACAGAGGCTGACCTCAGTGGCCGAGAGACCAAG 288
QY 61 GluGlnAlaAsnAlaValSerLysAlaValSerSerValAsnThrValAlaThrLys 80
Db 289 GAGCAGGCGCAACCGCGTGTGAGCAGGCTGTGTAGCAGCTCAACATGTGGCCACCAAG 348
QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValAlaArgLysGluAspLeu 100
Db 349 ACCGTGAGAGAGCGCGAAGCATCGCGTCACTCCGGGGGTGGTCCCAAGAGAGACTTG 408
QY 101 ArgProSerAlaProGlnGlnGlnGlyGluAlaSerLysGluLysGlnGluValAlaGlu 120
Db 409 AGGCATCTGCCCCCAAGAGAGGGTGAAGCATCCAAAGAGAAAGAGAAAGTGGCAGAG 468
QY 121 GluAlaGlnSerGlyGlyAsp 127
Db 469 GAGGCCAAGTGGGGAGAC 489

RESULT 5
; Sequence 171, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 171
; LENGTH: 796
; TYPE: DNA
; ORGANISM: Homo sapiens

```

US-10-102-806-171

Alignment Scores:

Pred. No.: 2,87e-59 Length: 796  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 14 Gaps: 0

US-09-017-715A-2 (1-127) x US-10-102-806-171 (1-796)

QY 1 MetAspValPheLysGlyPheSerIleAlaLysGlyValValGlyValAlaGlu 20  
DB 109 ATGATGCTTTCAAGAGGCGCTTCTCCATCGCCAGAGAGGGGTGGTGGCGGTGGAA 168  
QY 21 LysThrLysGlnGlyValThrGluAlaAlaGluLysThrLysGlnGlyValMetCysVal 40  
DB 169 AAGACCAAGCAGGGGGTGCAGAGCAGCTGAGAAAGACCAAGAGGGGGTCAATGATGTG 228  
QY 41 GYAlaLysThrLysGluAsnValValGlnSerValThrSerValAlaGluLysThrLys 60  
DB 229 GGAGCCAAAGACCAAGAGATGTGTACAGACGTGACTCAGTGCCGAGAGACCAAG 288  
QY 61 GIuGlnAlaAsnAlaValSerLysAlaValLysSerValAsnThrValAlaThrLys 80  
DB 289 GAGCAGGCCAAGCCGCTGAGCAGGCTGTGTGAGCAGCGTCAACCTGTGGCCACCAAG 348  
QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValArgLysGluAsnLys 100  
DB 349 ACCGTGAGAGAGCGGAGAACATCGCGCTCACCTCCGGGGTGGTGGCAAGAGGACTTG 408  
QY 101 ArgProSerAlaProGlnGlnGlnGlyGluLysGluLysGluGluValAlaGlu 120  
DB 409 AGGCCATCTGCCCCCAAGAGAGGTGTGAGCATTCAAAGAAAGAGAGGTGGCAGG 468  
QY 121 GIuAlaGlnSerGlyGlyAsp 127  
DB 469 GAGGCCCAAGAGTGGGGGAGAC 489

RESULT 6

US-10-826-157-5  
; Sequence 5, Application US/10826157  
; Publication No. US20050064548A1  
; GENERAL INFORMATION:  
; APPLICANT: Lindquist, Susan L.  
; APPLICANT: Outeiro, Tiago  
; TITLE OF INVENTION: YEAST ECTOPOICALLY EXPRESSING ABNORMALLY  
; FILE REFERENCE: 17481-003001  
; CURRENT APPLICATION NUMBER: US/10/826,157  
; CURRENT FILING DATE: 2004-04-16  
; PRIOR APPLICATION NUMBER: US 60/472,317  
; PRIOR FILING DATE: 2003-05-20  
; PRIOR APPLICATION NUMBER: US 60/463,284  
; PRIOR FILING DATE: 2003-04-16  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 384  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-826-157-5

Alignment Scores:

Pred. No.: 6,95e-59 Length: 384  
Score: 595.00 Matches: 124  
Percent Similarity: 99.21% Conservative: 2  
Best Local Similarity: 97.64% Mismatches: 1  
Query Match: 97.54% Indels: 0  
DB: 19 Gaps: 0

US-09-017-715A-2 (1-127) x US-10-826-157-5 (1-384)

QY 1 MetAspValPheLysGlyPheSerIleAlaLysGlyValValGlyValAlaGlu 20  
DB 1 ATGATGCTTTCAAGAGGCGCTTCTCCATCGCCAGAGAGGGGTGGTGGCGGTGGAA 60  
QY 21 LysThrLysGlnGlyValThrGluAlaAlaGluLysThrLysGlnGlyValMetCysVal 40  
DB 61 AAGACCAAGCAGGGGGTGCAGAGCAGCTGAGAAAGACCAAGAGGGGGTCAATGATGTG 120  
QY 41 GYAlaLysThrLysGluAsnValValGlnSerValThrSerValAlaGluLysThrLys 60  
DB 121 GGAGCCAAAGACCAAGAGATGTGTACAGACGTGACTCAGTGCCGAGAGACCAAG 180  
QY 61 GIuGlnAlaAsnAlaValSerLysAlaValLysSerValAsnThrValAlaThrLys 80  
DB 181 GAGCAGGCCAAGCCGCTGAGCAGGCTGTGTGAGCAGCGTCAACCTGTGGCCACCAAG 240  
QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValArgLysGluAsnLys 100  
DB 241 ACCGTGAGAGAGCGGAGAACATCGCGGTCACTCCGGGGTGGTGGCAAGAGGACTTG 300  
QY 101 ArgProSerAlaProGlnGlnGlnGlyGluLysGluLysGluGluValAlaGlu 120  
DB 301 AGGCCATCTGCCCCCAAGAGAGGTGTGAGCATTCAAAGAAAGAGAGGTGGCAGG 360  
QY 121 GIuAlaGlnSerGlyGlyAsp 127  
DB 361 GAGGCCCAAGAGTGGGGGAGAC 381

RESULT 7

US-10-097-340-297  
; Sequence 297, Application US/10097340  
; Publication No. US20030087250A1  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN  
; APPLICANT: Manjula GANNAVAPU  
; APPLICANT: Sebastian HOERSCH  
; APPLICANT: Shubhangi KAWATKAR  
; APPLICANT: Steve G. KOVATS  
; APPLICANT: Rachel E. MEYERS  
; APPLICANT: Michael MORRISSEY  
; APPLICANT: Peter OLANDT  
; APPLICANT: Ami SEN  
; APPLICANT: Peter VEIBY  
; APPLICANT: Gordon B. MILLS  
; APPLICANT: Robert C. BAST, Jr.  
; APPLICANT: Karen LU  
; APPLICANT: Rosemarie SCHMANDT  
; APPLICANT: Xumei ZHAO  
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
; FILE REFERENCE: Assessment, Prevention, and Therapy of Ovarian Cancer  
; CURRENT APPLICATION NUMBER: US/10/097,340  
; CURRENT FILING DATE: 2002-03-14  
; PRIOR APPLICATION NUMBER: 60/276,025  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/325,149  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/276,026  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/324,967  
; PRIOR FILING DATE: 2001/09/26  
; PRIOR APPLICATION NUMBER: 60/311,732  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/325,102  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/323,580  
; PRIOR FILING DATE: 2001-09-19  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 297  
; LENGTH: 720

```

: PRIOR APPLICATION NUMBER: US 60/368,919
: PRIOR FILING DATE: 2002-03-28
: NUMBER OF SEQ ID NOS: 564
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 469
: LENGTH: 720
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: allele
: LOCATION: 30,57,85,243,250,377,512,531,555,561,672
: OTHER INFORMATION: N is any
: US-10-282-174-469

Alignment Scores:
Pred. No.: 3,5e-58 Length: 720
Score: 592.00 Matches: 124
Percent Similarity: 97.64% Conservative: 0
Best Local Similarity: 97.64% Mismatches: 3
Query Match: 97.05% Indels: 0
DB: 17 Gaps: 0

US-09-017-715A-2 (1-127) x US-10-282-174-469 (1-720)

QY 1 MetAspValPheIyLysGlyPheSerIleAlaIyLysGlyValAlaGlyAlaValGlu 20
DB 49 ATGATGATGTTTCACAAAGAGGCGTTCTCCATCCGCACAAAGAGGCGGTGGTGGCGAA 108
QY 21 LysThrLysGlnGlyValIThrGluAlaAlaGlyLysThrLysGlnGlyValIThrVal 40
DB 109 AAGACCAAGCAGGGGGGTGACCGAACACCTGAGAAACCAAGAGGGGGTCAATGATGTG 168
QY 41 GlyAlaLysThrLysGluAsnValValGlnSerValIThrSerValAlaGlnLysThrLys 60
DB 169 GAGCCAAAGACCAAGAGAAATGTTTACAGACCGTGACTCACTGTCGCGAAGAACCAAG 228
QY 61 GluGlnAlaAsnAlaValSerLysAlaValIserSerValAsnThrValAlaThrLys 80
DB 229 GAGCAGCGCAACGCGGTGATGACNAGCGTGTGGTACAGCGTCAACACTGTGGCCACCAAG 288
QY 81 ThrValGluGluAlaGluAsnIleAlaValIThrSerGlyValAlaArgLysGluAspLeu 100
DB 289 ACCGTGAGGAGGCGAGAACATCCGGTCACTCCGGGGTGTGTGCGCAAGAGGACTTG 348
QY 101 ArgProSerAlaProGlnGlnGlnGlyValIserLysGluLysGlnGluValIAsn 120
DB 349 AGGCCATCTGCCCCCAACAGAGGGTNGGCATCAAGAAAGAGAAAGTGGCGAAG 408
QY 121 GluAlaGlnSerGlyLysAsp 127
DB 409 GAGGCCCAAGACTGGGGCGAGAC 429

RESULT 9
US-10-600-009-469
: Sequence 469, Application US/10600009
: Publication No. US2005009031A1
: GENERAL INFORMATION:
: APPLICANT: Becker, Kenneth David
: APPLICANT: Velicelebi, Gonul
: APPLICANT: Elliott, Kathryn J.
: APPLICANT: Wang, Xin
: APPLICANT: Tanzi, Rudolph E.
: APPLICANT: Berttram, Lars
: APPLICANT: Saunders, Aleister J.
: APPLICANT: Mullin, Kristina M.
: APPLICANT: Sampson, Andrew Johnson
: APPLICANT: Blacker, Deborah Lynne
: TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
: TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
: FILE REFERENCE: 37481-3308B
: CURRENT FILING DATE: 2003-06-18

```

PRIOR APPLICATION NUMBER: US 60/339,525  
PRIOR FILING DATE: 2001-10-25  
PRIOR APPLICATION NUMBER: US 60/338,010  
PRIOR FILING DATE: 2001-11-08  
PRIOR APPLICATION NUMBER: US 60/336,929  
PRIOR FILING DATE: 2001-11-08  
PRIOR APPLICATION NUMBER: US 60/338,363  
PRIOR FILING DATE: 2001-11-09  
PRIOR APPLICATION NUMBER: US 60/337,052  
PRIOR FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: US 60/368,919  
PRIOR FILING DATE: 2002-03-28  
PRIOR APPLICATION NUMBER: US 10/282,174  
PRIOR FILING DATE: 2002-10-25  
NUMBER OF SEQ ID NOS: 564  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 469  
LENGTH: 720  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: allele  
LOCATION: 30,57,85,243,250,377,512,531,555,561,672  
OTHER INFORMATION: N is any  
US-10-600-009-469

Alignment Scores:  
Pred. No.: 3.5e-58 Length: 720  
Score: 592.00 Matches: 124  
Percent Similarity: 97.64% Conservative: 0  
Best Local Similarity: 97.64% Mismatches: 3  
Query Match: 97.05% Indels: 0  
Gaps: 0  
DB: 19

US-09-017-715A-2 (1-127) x US-10-600-009-469 (1-720)

QY 1 MetAspValPheLysGlyPheSerIleAlaLysGlyValValGluValAlaGlu 20  
DB 49 ATGAGATGTTTCAAGAGGGCTTCTCATCGCCAGAGGGCGGTGCGGTGCGA 108  
QY 21 LysThrLysGlnGlyValThrGluAlaAlaGluLysThrLysGluGlyValMetLysVal 40  
DB 109 AAGACCAAGCAGGGGTGACGAGACAGCTGAGAACACAGAGGGGTCTGATATGTG 168  
QY 41 GAlaLysThrLysGluAsnValValGlnSerValThrSerValAlaGluLysThrLys 60  
DB 169 GAGCCAGACCAAGAGAAATGTTGACAGCGGTCACTCGAGCCAGAAAGACCAAG 228  
QY 61 GlnGlnAlaAsnAlaValSerLysAlaValAlaSerSerValAsnThrValAlaThrLys 80  
DB 229 GAGCAGGCCACGACGNGTGAAGAGGCTGTGTGAGCAGCGTCAACACTGTGGCCACCAAG 288  
QY 81 ThrValGluGluValAlaGluAsnIleAlaValThrSerGlyValValArgLysGluAspLeu 100  
DB 289 ACCGTGAGAGAGCGGAGAAATCGCGTCACTCGGAGGTGGTGGCAAGAGGACTTG 348  
QY 101 AcGProSerAlaProGlnGlnGluGluValAlaSerLysGluLysGluGluValAlaGlu 120  
DB 349 AGGCACTCTGCCCCCAAGAGGGGTGNGGCTCCAAAGAGAAAGAGAGTGGCAGAG 408  
QY 121 GluAlaGlnSerGlyGlyAsp 127  
DB 409 GAGGCCCAAGAGTGGGGAAC 429

RESULT 10

US-09-918-995-2705  
Sequence 2705, Application US/09918995  
Publication No. US20030073623A1  
GENERAL INFORMATION:  
APPLICANT: HySeq, Inc.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
FROM VARIOUS CDNA LIBRARIES  
FILE REFERENCE: 20411-756

CURRENT APPLICATION NUMBER: US/09/918,995  
CURRENT FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: US/09/235,076  
PRIOR FILING DATE: 1999-01-20  
NUMBER OF SEQ ID NOS: 38054  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2705  
LENGTH: 479  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(479)  
OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-2705

Alignment Scores:  
Pred. No.: 3.03e-52 Length: 479  
Score: 538.00 Matches: 112  
Percent Similarity: 99.14% Conservative: 3  
Best Local Similarity: 96.55% Mismatches: 1  
Query Match: 98.20% Indels: 0  
Gaps: 0  
DB: 10

US-09-017-715A-2 (1-127) x US-09-918-995-2705 (1-479)

QY 12 LysLysGlyValValGluValAlaGluLysThrLysGlnGlyValThrGluAlaGlu 31  
DB 47 CAGAGGGCGGTGAGNGGTCGCGTGGAAGACCAAGCAGGGGTGACGGAAGCAGCTGAG 106  
QY 32 LysThrLysGluGlyValMetLysValGluAlaLysThrLysGluLysValAlaGlnSer 51  
DB 107 AAGACCAAGAGGGGTCTGATGTGAGGACCAAGACCAAGAGAAATGTTCTACAGAGC 166  
QY 52 ValThrSerValAlaGluLysThrLysGlnGlnAlaAsnAlaValSerLysAlaValAla 71  
DB 167 GTGACCTCAGTGGCCGAGAGACCAAGAGCAGGCCAAGCGGTGAGAGGCTGTGGTG 226  
QY 72 SerSerValAsnThrValAlaThrLysThrValGluGluAlaGluAsnIleAlaValThr 91  
DB 227 AGCAGCGTCAACACTGTGGCCCAAGACCGTGGAGAGGGCGGAGAACATCGCGGTCAAC 286  
QY 92 SerGlyValValArgLysGluAspLeuArgProSerAlaProGlnGlnGluGluValAla 111  
DB 287 TCCGGGGTGTGTCGCAAGAGGACTTGAGGCCATCTGCCCCCAAGAGAGGTGAGCA 346  
QY 112 SerLysGluLysGluGluValAlaGluGluValAlaGlnSerGlyGlyAsp 127  
DB 347 TCCAAAGAGAAAGAGAGTGGCAGAGAGGCCCAAGTGGGGAGAC 394

RESULT 11

US-10-267-849-1  
Sequence 1, Application US/10267849  
Publication No. US20030087824A1  
GENERAL INFORMATION:  
APPLICANT: Ji, Hongjun  
TITLE OF INVENTION: Breast Cancer Specific Gene 2  
FILE REFERENCE: 1488.0810001  
CURRENT APPLICATION NUMBER: US/10/267,849  
CURRENT FILING DATE: 2002-10-10  
PRIOR APPLICATION NUMBER: US/08/673,284  
PRIOR FILING DATE: 1996-06-28  
PRIOR APPLICATION NUMBER: US 60/000,602  
PRIOR FILING DATE: 1995-06-30  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 786  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-267-849-1



Alignment Scores:

Pred. No.: 5.09e-44 Length: 786  
 Score: 468.50 Matches: 111  
 Percent Similarity: 86.15% Conservative: 1  
 Best Local Similarity: 85.38% Mismatches: 15  
 Query Match: 76.80% Indels: 4  
 DB: 14 Gaps: 1

US-09-017-715A-2 (1-127) x US-10-267-849-1 (1-786)

Qy 1 MetaspValpheylysglypheyserilealalyserglyvalalglalavalglu 20  
 Db 95 ATGGATGTCTTCAAGAGGGCTTCTCCATCGCCAAAGAGGGCGTGGCGGTGGAA 154  
 Qy 21 LysThrlysglncglvalalthrglualalaglulysThrlysglulglvalmetTyVal 40  
 Db 155 AAGGCCAAGAGAGGGGGTGGAGAGAGCTGAGAGAGCAAGAGGGGGTCAATATATGTG 214  
 Qy 41 G1YAlalyserThrlysgluAan---Val-ValGlnSerValThrservalalaglulysTh 59  
 Db 215 GAGGCCAAGACCAAGAGAAATGTTGTATGTCAGAGCTGACCTCACTGCCGAGAAAGAC 274  
 Qy 59 rlysglulglualaanaalavalserlysalavalserervalanthrvalalath 79  
 Db 275 CAAGAGCAGAGCCCAAGCCCGTGAAGCAAGCGTGTGAGAGAGCTCAACACTKTGGCCAC 334  
 Qy 79 rlyserThrValGlulglualaglualenilealavalThrservlValalarglysgluas 99  
 Db 335 CAAGACCGTGAAGAGCGGAGAAATCGCGTCACTCCGGGTGTCGCGCAAGAGAGA 394  
 Qy 99 pleuarProseralProglnglncglulglualaserlysglu-LyeglulgluVala 119  
 Db 395 YTKKAGGCCATY-TRKCCCAACAGAGCGGTGAGGCGATCARAGARARAKWGAAGWG 453  
 Qy 119 lacluglualaglnserglylasp 127  
 Db 454 CMRAKKRGKMSGAGAGTGGGAGAGAC 479

RESULT 12

US-10-204-337A-5  
 ; Sequence 5, Application US/10204337A  
 ; Publication No. US20040128706A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Maeliah, Eliezer  
 ; TITLE OF INVENTION: Method for screening for Anti-Amyloidogenic Properties and Method  
 ; FILE REFERENCE: 6627-PC9014  
 ; CURRENT APPLICATION NUMBER: US/10/204,337A  
 ; PRIOR FILING DATE: 2002-08-16  
 ; PRIOR APPLICATION NUMBER: US 60/183,571  
 ; PRIOR FILING DATE: 2000-02-18  
 ; PRIOR APPLICATION NUMBER: PCT/US00/07216  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 210  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-204-337A-5

Alignment Scores:

Pred. No.: 1.02e-28 Length: 210  
 Score: 328.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 2  
 Best Local Similarity: 97.14% Mismatches: 0  
 Query Match: 53.77% Indels: 0  
 DB: 18 Gaps: 0

US-09-017-715A-2 (1-127) x US-10-204-337A-5 (1-210)

Qy 1 MetaspValpheylysglypheyserilealalyserglyvalalglalavalglu 20  
 Db 28 ATGGATGTCTTCAAGAGGGCTTCTCCATCGCCAAAGAGGGCGTGGCGGTGGAA 87

Db 1 ATGGATGTCTTCAAGAGGGCTTCTCCATCGCCAAAGAGGGCGTGGCGGTGGAA 60

Qy 21 LysThrlysglncglvalalthrglualalaglulysThrlysglulglvalmetTyVal 40  
 Db 61 AAGGCCAAGAGAGGGGGTGGAGAGAGCTGAGAGAGCAAGAGGGGGTCAATATATGTG 120

Qy 41 G1YAlalyserThrlysgluAanvalalglnservalThrservalalaglulysThrly 60  
 Db 121 GAGGCCAAGACCAAGAGAAATGTTGTATGTCAGAGCTGACCTCACTGCCGAGAAAGCAAG 180

Qy 61 Glulglualaanaalavalserlysalavalserervalanthrvalalath 70  
 Db 181 GAGCAGGCCAAGCGGTGAGAGAGCTGTG 210

RESULT 13

US-10-152-319A-1710  
 ; Sequence 1710, Application US/10152319A  
 ; Publication No. US20040072160A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mendrick, Donna  
 ; APPLICANT: Porter, Mark  
 ; APPLICANT: Johnson, Kory  
 ; APPLICANT: H399, Brandon  
 ; APPLICANT: Castle, Arthur  
 ; APPLICANT: Blaschoff, Michael  
 ; TITLE OF INVENTION: Molecular Toxicology Modeling  
 ; FILE REFERENCE: 44921-5089-US  
 ; CURRENT APPLICATION NUMBER: US/10/152,319A  
 ; PRIOR FILING DATE: 2002-05-22  
 ; PRIOR APPLICATION NUMBER: US 60/292,335  
 ; PRIOR FILING DATE: 2001-05-22  
 ; PRIOR APPLICATION NUMBER: US 60/297,523  
 ; PRIOR FILING DATE: 2001-06-13  
 ; PRIOR APPLICATION NUMBER: US 60/298,925  
 ; PRIOR FILING DATE: 2001-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/303,810  
 ; PRIOR FILING DATE: 2001-07-10  
 ; PRIOR APPLICATION NUMBER: US 60/303,807  
 ; PRIOR FILING DATE: 2001-07-10  
 ; PRIOR APPLICATION NUMBER: US 60/303,808  
 ; PRIOR FILING DATE: 2001-07-10  
 ; PRIOR APPLICATION NUMBER: US 60/315,047  
 ; PRIOR FILING DATE: 2001-08-28  
 ; PRIOR APPLICATION NUMBER: US 60/324,928  
 ; PRIOR FILING DATE: 2001-09-27  
 ; PRIOR APPLICATION NUMBER: US 60/330,867  
 ; PRIOR FILING DATE: 2001-11-01  
 ; PRIOR APPLICATION NUMBER: US 60/330,462  
 ; PRIOR FILING DATE: 2001-10-22  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 2221  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1710  
 ; LENGTH: 1018  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. NM\_019169  
 ; US-10-152-319A-1710

Alignment Scores:

Pred. No.: 1.91e-26 Length: 1018  
 Score: 316.00 Matches: 70  
 Percent Similarity: 68.60% Conservative: 13  
 Best Local Similarity: 57.85% Mismatches: 32  
 Query Match: 51.80% Indels: 6  
 DB: 17 Gaps: 1

US-09-017-715A-2 (1-127) x US-10-152-319A-1710 (1-1018)

Qy 1 MetaspValpheylysglypheyserilealalyserglyvalalglalavalglu 20  
 Db 28 ATGGATGTCTTCAAGAGGGCTTCTCCATCGCCAAAGAGGGCGTGGCGGTGGAA 87

Oy	2	LYSTHLYGSLGSLVALThrGluAlaAlaGluLSTHLYSGluGlyValMetLysVal	40
Db	88	AAACCAACGAGGGTGTGGCAGAGCAGCTGGGAAAGACAAAGAAGGGGTCTTATGTA	147
Oy	41	GLVALAlaLysThrLysGluSuaValValGlnSerValLThrSerValAlaGluLysThrLys	60
Db	148	GGTTCCAAACCTACGAGAGGAGTGGTTCATGTGAGTGAACAACGTGGCTGACAAACCAA	207
Oy	61	GLGluAlaAlaSuaValValSerLysValValLysSerSerValAsnThrValAlaThrLys	80
Db	208	GACCAACTGCACAAATTGTGGAGGGCAGCGTGAATGGTGTGACACGACTGGCTCAGAG	267
Oy	81	ThrValGluGluValaGluAsnLLeaValLThrSerGlyValValArgLysGluAspLeu	100
Db	268	ACAGTGAAGGAGACTGTGAACAATTGCTGTGCACCTGGTTTGTCCAAAGACCAAGATG	327
Oy	101	ArgProSerAlaProGluGlnGluGlyValaLysLysGluLysGluGluValaGlu	120
Db	328	-----GGCAAGGTGAAGAAGGATCCACCAAGAGGAAATCTCGAA	365

```

RESULT 14
US-10-486-706-260
: Sequence 260, Application US/10486706
: Publication No. US20050071086A1
: GENERAL INFORMATION:
: APPLICANT: LANDEFIELD, PHILIP W.
: APPLICANT: BLALOCK, ERIC M.
: APPLICANT: CHEN, KUEY-CHU
: APPLICANT: FOSTER, THOMAS C.
: TITLE OF INVENTION: GENE EXPRESSION PROFILE BIOMARKERS AND THERAPEUTIC TARGETS FOR
: TITLE OF INVENTION: BRAIN AGING AND AGE-RELATED COGNITIVE IMPAIRMENT
: FILE REFERENCE: 50229-426
: CURRENT APPLICATION NUMBER: US/10/486,706
: PRIOR FILING DATE: 2004-02-13
: PRIOR APPLICATION NUMBER: PCT/US02/25667
: PRIOR FILING DATE: 2002-08-13
: PRIOR APPLICATION NUMBER: US 60/311,343
: PRIOR FILING DATE: 2001-08-13
: NUMBER OF SEQ ID NOS: 461
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 260
: LENGTH: 1018
: TYPE: DNA
: ORGANISM: Rattus norvegicus
US-10-486-706-260

```

Alignment Scores:	
Pred. No.:	1,91e-26
Length:	1018
Score:	316.00
Percent Similarity:	68.60%
Best Local Similarity:	57.85%
Query Match:	51.80%
DB:	15
Gaps:	1

Qy	1	Me	Asp	Val	Phe	ly	S	G	I	Phe	Ser	I	Le	Al	Val	S	G	I	Val	Al	Val	Al	G	I	u	20	
Db	28	A	T	G	A	T	G	T	T	C	A	T	T	T	A	A	G	G	C	A	A	G	A	G	G	T	87
Qy	21	L	y	S	T	H	r	i	S	G	I	n	G	I	Val	I	n	H	r	G	I	A	a	I	a	G	40
Db	88	A	A	A	C	C	A	G	A	G	G	G	T	G	T	G	C	A	G	G	A	G	A	C	A	A	147
Qy	41	G	I	Val	A	L	S	T	H	r	i	S	G	I	u	S	n	Val	I	G	I	n	S	er	Val	I	60
Db	148	G	G	T	T	C	A	A	A	C	T	A	A	C	T	A	G	A	G	G	A	G	A	C	T	G	207

[illegible]

```

RESULT 15
US-10-737-262-2
; Sequence 2, Application US/10737262
; Publication No. US20040197315A1
; GENERAL INFORMATION:
; APPLICANT: Stefanis, Leonidas
; APPLICANT: Greene, Lloyd A.
; TITLE OF INVENTION: Dopaminergic Cell Lines Stably Expressing A53T Alpha-Synuclein
; TITLE OF INVENTION: and Methods of Using Same
; FILE REFERENCE: 5199-26
; CURRENT APPLICATION NUMBER: US/10/737,262
; CURRENT FILING DATE: 2003-12-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-737-262-2

```

Alignment Scores:	
Pred. No.:	1,58e-26
Score:	312.50
Percent Similarity:	71.82%
Best Local Similarity:	62.73%
Query Match:	51.23%
DB:	16
US-09-017-715A-2 (1-127) × US-10-737-262-2 (1-437)	
	Length: 437
	Matches: 69
	Conservative: 10
	Mismatches: 28
	Indels: 3
	Gaps: 1

Qy	MeAspVal.PheIysIysGlyPheSerIleAlaIysGlyValAlaValAlu	20
Db	ATGATGATTCATGAAAGACCTTCGAAAGCCAGAGGAGTTGGCTGCCTGAG	74
Qy	LyeThrLysGlnGlyValThrGluAlaIleGluIysThrLysGlnGlyValMetTyrVal	40
Db	AAAAACCAACAGGGGTGTGGCAAAACACAGAAAGCAAAAGAGGGGTCTCTATGTA	134
Qy	GlyValAlaIysThrLysGluAsnValValGlnSerValThrSerValAlaGluIysThrLys	60
Db	GGCTCCAAAACCAAGGAGGAGGTGTGTCATGTGTGACAAACAGTGGCTGAGAAACCAA	194
Qy	GluGlnAlaAsnAlaValSerIysValAlaValSerSerValAsnThrValAlaThrLys	80
Db	GACCAAGTACAAATGTTGGAGACACAGTGGTGAACGGGTGTGACGACGATGCCAGAG	254
Qy	ThrValGluGluAlaGluIleAlaValThrSerGlyValValArgLysGluAspLeu	100
Db	ACAGTGGAGGAGACGAGGACATTGCAGAGCAACTGCTTTGTCAAAAAGACCAAGTTG	314
Qy	ArgProSerIleProGlnGln	107
Db	GGCAAGATGAAAGAGACCCCAACAGAA	344

Search completed: May 4, 2005, 16:39:31  
Job time : 1042.71 secs

THIS PAGE LEFT BLANK

---

**THIS PAGE LEFT BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - nucleic search, using frame\_p2n model

Run on: May 4, 2005, 09:07:56 ; Search time 5778.07 Seconds

(without alignments)  
836.639 Million cell updates/sec

Title: US-09-017-715A-2

Perfect score: 610  
Sequence: 1 MDVFKKGFSJAKKGVGAVE.....EGEASKEKEVAERAGSGD 127

## Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delpext 7.0	

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68473088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Command line parameters:

-MODEL=frame\_p2n.model -DEV=x1h  
-Q=/cgn2\_1/USPRO.spool\_h/US09017715/runat\_04052005\_100744\_25619/app\_query.fasta\_1.661  
-DB=EST -QPM=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsm62 -TRANS=human40.cdi -LIST=45  
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09017715\_@CGN\_1\_1\_5334\_@runat\_04052005\_100744\_25619 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEW TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

EST: \*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_ges1.\*  
9: gb\_ges2.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	602	98.7	555	7 CV028548 7090 Full1
2	602	98.7	578	5 BP212912 BP212912
3	602	98.7	582	5 BP197662 BP197662
4	602	98.7	584	5 BP201686 BP201686
5	602	98.7	653	4 B1757131 603030894
6	602	98.7	751	4 B1836596 603089575
7	602	98.7	781	5 BQ893395 AGENCOURT
8	602	98.7	809	5 BM921124 AGENCOURT
9	602	98.7	855	5 BQ882072 AGENCOURT

10	602	98.7	884	5 BQ439430
11	602	98.7	891	5 BQ221776
12	602	98.7	903	4 BG286466
13	595	97.5	538	4 B1548891
14	595	97.5	568	4 BG708703
15	595	97.5	617	4 BG707764
16	595	97.5	659	7 CN410061
17	595	97.5	755	4 B1597796
18	595	97.5	824	4 B1600882
19	591	96.9	765	4 B1603159
20	591	96.9	768	4 B1603171
21	589	96.6	937	5 BUI57619
22	589	96.6	1009	5 BUI79779
23	576	94.4	641	5 AL712443
24	575	94.3	555	5 BX474500
25	567	93.0	623	5 BP381244
26	566.5	92.9	799	4 B1488930
27	564.5	92.5	949	5 BQ068800
28	557	91.3	706	4 BG328738
29	555	91.0	583	5 BP200612
30	548	89.8	583	5 BP346497
31	542	88.9	756	5 BQ901053
32	541	88.7	489	5 BP201709
33	526	86.2	462	5 BX474511
34	524	85.9	467	6 CB107161
35	523	85.7	625	2 AM659211
36	523	85.7	625	7 CR452251
37	523	85.7	625	7 CR454933
38	523	85.7	646	7 CR551748
39	523	85.7	670	7 CO879176
40	523	85.7	689	7 CR454179
41	523	85.7	700	7 CK770094
42	523	85.7	801	7 CO873848
43	522.5	85.7	721	7 CN157077
44	522.5	85.7	721	7 CN159030
45	519	85.1	799	7 CO877016

## ALIGNMENTS

RESULT 1  
CV028548  
LOCUS  
DEFINITION 7090 Full Length cDNA from the Mammalian Gene Collection Homo  
sapiens CDNA 5' similar to BC014098, mRNA sequence.  
CV028548  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 555)  
Rual J.F., Hirozane-Kishikawa T., Hao T., Bertin N., Li S.,  
Clingingsmith T.R., Hartley J.L., Bosio D., Cheo D., Moore T.,  
Simmons B., Sequerra R., Bosak S., Doucette-Stamm L., Le Pench C.,  
Vandenhaute J., Clusick M.E., Alcala J.S., Hill D.E. and Vidal M.  
Human ORFeome Version 1.1: a Platform for Reverse Proteomics  
Genome Res. (2004) In press  
Contact: Vidal M

TITLE JOURNAL  
COMMENT  
Marc Vidal Laboratory  
Dana Farber Cancer Institute  
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA  
Tel: 617 632 5180  
Fax: 617 632 5739  
Email: Marc.Vidal@dfci.harvard.edu  
ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF  
results from a PCR reaction using an MGC full-length cDNA as  
template DNA and ORF specific primers  
PCR Primers  
FORWARD: ATGATGTCTTCAAGAAGGCTTCTC  
BACKWARD: TAGTCTCCCACTCTGG

Insert Length: 555 Std Error: 48.00  
 Seq: 11017 row: 05 column: B  
 Read primer: ACTGCCCTCCTTACACGTCGTACTGGAAGAAC  
 High quality sequence start: 97  
 High quality sequence stop: 554  
 POLYA=No.

FEATURES  
 source  
 Location/Qualifiers  
 1..555  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="mixed"  
 /clone\_lib="Full Length cDNA from the Mammalian Gene Collection"  
 /note="Vector: mixed; The ORFs were PCR amplified from the MGC (Mammalian Gene Collection) as of April 2004 and cloned by recombinational Gateway cloning into pDONR223 Donor vector. Reference : MGC (Mammalian Gene Collection) Program Team, Generation and Initial Analysis of more than 15,000 Full-length Human and Mouse cDNA Sequences. FNAS, 2002, 99(26), 16899-16903"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,39e-58 Length: 555  
 Score: 602.00 Matches: 125  
 Percent Similarity: 100.00% Conservative: 2  
 Best Local Similarity: 98.43% Mismatches: 0  
 Query Match: 98.69% Indels: 0  
 Gaps: 7 0

US-09-017-715a-2 (1-127) x CVO28548 (1-555)

QY 1 MetAepValPheIyLysGlyPheSerIleAlaIyLysGlyValIaGlyValaIaGlu 20  
 DB 1 ATGGAGTCTTCAAAAAGGGCTTCTCCATCGCCAAAGAGCGCTGGTGGGTGGGAA 60  
 QY 21 LysThrLysGlnGlyValThrGluAlaIaGluLysThrLysGlnGlyValMetYrVal 40  
 DB 61 AAGACCAAGCAGGGGGTGACGAGACAGCTGAGAAACCAAGAGGGGCTCATGTATG 120  
 QY 41 GlyAlaLysThrLysGlnAsnValIaGlnSerValThrSerValaIaGluLysThrLys 60  
 DB 121 GGAGCCAAAGCAGAGATGTTGACAGAGCGTCACTCACTGCGCAGAAAGCAAG 180  
 QY 61 GluGlnAlaAsnAlaValSerLysAlaValaIaSerSerValaAsnThrValaIaThrLys 80  
 DB 181 GAGCAGGCAACCGCCGTGAGCGAGCTGTGTAGACAGCTCAACACTGTGGCCACCAAG 240  
 QY 81 ThrValIGluGluAlaGluAsnIleAlaValaIaThrSerGlyValaIaArgLysGluAspLeu 100  
 DB 241 ACCGTGAGAGGAGCGGAGAACTCGCGGTCACTCCGGGGTGTGGCAAGAGAGCTTG 300  
 QY 101 ArgProSerAlaProGlnGlnGlnGlyGluAlaSerLysGlnLysGlnGluValaIaGlu 120  
 DB 301 AGGCCATCTGCCCGCCCAAGAGAGGTGTAGGCTCAAGAAAGAAAGAGAGAGTGGCAGAG 360  
 QY 121 GluAlaGlnSerGlyLysp 127  
 DB 361 GAGGCCCAAGTGGGAGAC 381

RESULT 2  
 LOCUS BP12912 578 bp mRNA linear EST 15-SEP-2004  
 DEFINITION BP12912 Sugano cDNA library, cerebrum Homo sapiens cDNA clone  
 CBR05118, mRNA sequence.  
 ACCESSION BP12912  
 VERSION BP12912.1 GI:52085803  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 578)  
 AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,  
 Mizushima-Sugano,J., Nakai,K. and Sugano,S.  
 TITLE Sequence comparison of human and mouse genes reveals a homologous  
 block structure in the promoter regions  
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)  
 COMMENT Contact: Yutaka Suzuki  
 Department of Virology  
 Institute of Medical Science, University of Tokyo  
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
 Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES  
 source  
 Location/Qualifiers  
 1..578  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CBR05118"  
 /tissue\_type="cerebrum"  
 /clone\_lib="Sugano cDNA library, cerebrum"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,47e-58 Length: 578  
 Score: 602.00 Matches: 125  
 Percent Similarity: 100.00% Conservative: 2  
 Best Local Similarity: 98.43% Mismatches: 0  
 Query Match: 98.69% Indels: 0  
 Gaps: 5 0

US-09-017-715a-2 (1-127) x BP12912 (1-578)

QY 1 MetAepValPheIyLysGlyPheSerIleAlaIyLysGlyValIaGlyValaIaGlu 20  
 DB 125 ATGGAGTCTTCAAAAAGGGCTTCTCCATCGCCAAAGAGCGCTGGTGGGTGGGAA 184  
 QY 21 LysThrLysGlnGlyValThrGluAlaIaGluLysThrLysGlnGlyValMetYrVal 40  
 DB 185 AAGACCAAGCAGGGGGTGACGAGACAGCTGAGAAACCAAGAGGGGCTCATGTATG 244  
 QY 41 GlyAlaLysThrLysGlnAsnValIaGlnSerValThrSerValaIaGluLysThrLys 60  
 DB 245 GGAGCCAAAGCAGAGATGTTGACAGAGCGTCACTCACTGCGCAGAAAGCAAG 304  
 QY 61 GluGlnAlaAsnAlaValSerLysAlaValaIaSerSerValaAsnThrValaIaThrLys 80  
 DB 305 GAGCAGGCAACCGCCGTGAGCGAGCTGTGTAGACAGCTCAACACTGTGGCCACCAAG 364  
 QY 81 ThrValIGluGluAlaGluAsnIleAlaValaIaThrSerGlyValaIaArgLysGluAspLeu 100  
 DB 365 ACCGTGAGAGGAGCGGAGAACTCGCGGTCACTCCGGGGTGTGGCAAGAGAGCTTG 424  
 QY 101 ArgProSerAlaProGlnGlnGlnGlyGluAlaSerLysGlnLysGlnGluValaIaGlu 120  
 DB 425 AGGCCATCTGCCCGCCCAAGAGAGGTGTAGGCTCAAGAAAGAAAGAGAGTGGCAGAG 484  
 QY 121 GluAlaGlnSerGlyLysp 127  
 DB 485 GAGGCCCAAGTGGGAGAC 505

RESULT 3  
 LOCUS BP197662 582 bp mRNA linear EST 14-SEP-2004  
 DEFINITION BP197662 Sugano cDNA library, adrenal gland Homo sapiens cDNA clone  
 ADG06551, mRNA sequence.  
 ACCESSION BP197662  
 VERSION BP197662.1 GI:52043849  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 582)  
 AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,

TITLE Mizushima-Sugano,J., Nakai,K. and Sugano,S.  
Sequence comparison of human and mouse genes reveals a homologous  
block structure in the promoter regions  
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)  
COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES  
source  
1. .582  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="ADG06551"  
/tissue\_type="adrenal gland"  
/clone\_id="Sugano cDNA library, adrenal gland"

ORIGIN

Alignment Scores:  
Pred. No.: 1,486-58 Length: 582  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 5 Gaps: 0

US-09-017-715a-2 (1-127) x BP197662 (1-582)

Qy 1 MetApyAlPheLySgLyPheSerIleAlaLySgLyValValGlyAlaValGlu 20  
Db 85 ATGATGTCCTTCAAGAGGCGCTTCTCATGCCAAGAGGCGGTGGCGGTGGA 144  
Qy 21 LyThrLySgInGlyValThrGluAlaGluLyThrLySgInGlyValMetTyVal 40  
Db 145 AAGACCAAGAGGGGGTGACGGAAGCAGCTGAGAAAGCAAGAGGGGGTCATATGTG 204  
Qy 41 GlyAlaLyThrLySgInGluValGlnSerValThrSerValAlaGluLyThrLyS 60  
Db 205 GAGGCCAAGACCAAGAGAAATGTTGACAGCGCTGACCTGAGCGCGAAGACCAAG 264  
Qy 61 GluGlnAlaAsnAlaValSerLySAlaValAlaSerSerValAanthrValAlaThrLyS 80  
Db 265 GAGCAGGCCAAGCGCGTGAAGGCTGTGTGACGCGTCAACACTGTGGCCACCAAG 324  
Qy 81 ThrValGluGluAlaGluAenIleAlaValThrSerGlyValValAArgLySgInGluAspLeu 100  
Db 325 ACCGTGAGAGGCGGAGAACATCGCGTCACTCCGGGGTGTGCGCAAGAGGACTTG 384  
Qy 101 ArgProSerAlaProGlnGlnGluGlyGluAlaSerLySgInGluGluValAlaGlu 120  
Db 385 AGGCCATCTGCCCGCCCAAGAGGGTGAAGCATCCAAAGAAAGAGAAAGTGGCAGAG 444  
Qy 121 GluAlaGlnSerGlyGlyAsp 127  
Db 445 GAGGCCAGAGTGGGGAGAC 465

RESULT 4  
BP201686 584 bp mRNA linear EST 14-SEP-2004  
LOCUS BP201686 Sugano cDNA library, amygdala Homo sapiens cDNA clone  
DEFINITION AMR06165, mRNA sequence.  
ACCESSION BP201686  
VERSION BP201686.1 GI:52051909  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 584)  
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,  
Mizushima-Sugano,J., Nakai,K. and Sugano,S.  
TITLE Sequence comparison of human and mouse genes reveals a homologous

JOURNAL block structure in the promoter regions  
COMMENT Genome Res. 14 (9), 1711-1718 (2004)  
Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES  
source  
1. .584  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="AMR06165"  
/tissue\_type="amygdala"  
/clone\_id="Sugano cDNA library, amygdala"

ORIGIN

Alignment Scores:  
Pred. No.: 1,496-58 Length: 584  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 5 Gaps: 0

US-09-017-715a-2 (1-127) x BP201686 (1-584)

Qy 1 MetApyAlPheLySgLyPheSerIleAlaLySgLyValValGlyAlaValGlu 20  
Db 125 ATGATGTCCTTCAAGAGGCGCTTCTCATGCCAAGAGGCGGTGGCGGTGGA 184  
Qy 21 LyThrLySgInGlyValThrGluAlaGluLyThrLySgInGlyValMetTyVal 40  
Db 185 AAGACCAAGAGGGGGTGACGGAAGCAGCTGAGAAAGCAAGAGGGGGTCATATGTG 244  
Qy 41 GlyAlaLyThrLySgInGluValGlnSerValThrSerValAlaGluLyThrLyS 60  
Db 245 GAGGCCAAGACCAAGAGAAATGTTGACAGCGCTGACCTGAGCGCGAAGACCAAG 304  
Qy 61 GluGlnAlaAsnAlaValSerLySAlaValAlaSerSerValAanthrValAlaThrLyS 80  
Db 305 GAGCAGGCCAAGCGCGTGAAGGCTGTGTGACGCGTCAACACTGTGGCCACCAAG 364  
Qy 81 ThrValGluGluAlaGluAenIleAlaValThrSerGlyValValAArgLySgInGluAspLeu 100  
Db 365 ACCGTGAGAGGCGGAGAACATCGCGTCACTCCGGGGTGTGCGCAAGAGGACTTG 424  
Qy 101 ArgProSerAlaProGlnGlnGluGlyGluAlaSerLySgInGluGluValAlaGlu 120  
Db 425 AGGCCATCTGCCCGCCCAAGAGGGTGAAGCATCCAAAGAAAGAGAAAGTGGCAGAG 484  
Qy 121 GluAlaGlnSerGlyGlyAsp 127  
Db 485 GAGGCCAGAGTGGGGAGAC 505

RESULT 5  
B1757131 653 bp mRNA linear EST 25-SEP-2001  
LOCUS B1757131  
DEFINITION 603030894F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:520162 5',  
mRNA sequence.  
ACCESSION B1757131  
VERSION B1757131.1 GI:15748709  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Straube, Ph.D.  
COMMENT Email: cgabs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1M1503 row: 1 column: 23  
High quality sequence stop: 649.  
Location/Qualifiers

## FEATURES

source

1. 653  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5201062"  
/lab\_host="DH10B"  
/clone\_11b="NIH\_MGC\_114"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6  
male brains, age range 23-27 yo. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.5 kb, insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 019. Note:  
this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.72e-58 Length: 653  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2 (1-127) x B1757131 (1-653)

QY 1 MetAaPvAlPheLySgLyPheSerIleAlaLySgLyVAlVaIgLyAlaVaIgLu 20  
DB 57 ATGAGATCTTCAAGAGGGCTTCTCCATCGCCAGAGGGCGTGGTGGCGGAGAA 116  
QY 21 LySThrLySGInGlyValThrGluAlaAlaGluLyThLySGInGlyValMetTyVal 40  
DB 117 AAGACCAACAGGGGGTGAAGAGACGCTGAAGAACCAAGAGGGGGTCAATGATG 176  
QY 41 G1yAlaLyThLySGInGluSnVAlValGInSerValThSerValAlaGluLyThLyS 60  
DB 177 GGAGCCAAAGACAGAGATGTTGTACAGAGCGTCACTCACTGCGCAAGAACCAAG 236  
QY 61 GluGlnAlaAsnAlaValSerIleAlaValSerValAsnThrValAlaThrLyS 80  
DB 237 GAGCAGGCCAAGCGCTGAGCGGTGTGTAGAGCGTCAACCTGTGGCCACCAAG 296  
QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValArgLySGInPleu 100  
DB 297 ACCGTGAGAGAGCGGAGAACATCGCGTCACTCCGGGGTGTGGCAAGAGACTTG 356  
QY 101 ArgProSerAlaProGInGInGluGlyGluAlaSerLySGInGluGluValAlaGlu 120  
DB 357 AGGCCATCTGCCCCCAACAGAGGGGTGAGCATCCAAAGAGAAAGAGAGTGGCAGAG 416  
QY 121 GluAlaGlnSerGlyGlyAsp 127  
DB 417 GAGGCCCAAGTGGGGAGAGC 437

RESULT 6  
B1836596 751 bp mRNA linear EST 04-OCT-2001  
LOCUS 603089575F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5228538 5',  
DEFINITION mRNA sequence.  
ACCESSION B1836596  
VERSION B1836596.1 GI:15948146

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 751)  
TITLE NIH-MGC http://mhc.nci.nih.gov/  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1M1575 row: e column: 19  
High quality sequence stop: 703.  
Location/Qualifiers

## FEATURES

source

1. 751  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5228538"  
/lab\_host="DH10B"  
/clone\_11b="NIH\_MGC\_120"  
/note="Organ: pooled pancreas and spleen; Vector:  
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
source anonymous pool of spleen and pancreas from 28 yo  
male. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 1.5 kb, insert size range 1-2.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 025. Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 2.06e-58 Length: 751  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2 (1-127) x B1836596 (1-751)

QY 1 MetAaPvAlPheLySgLyPheSerIleAlaLySgLyVAlVaIgLyAlaVaIgLu 20  
DB 96 ATGAGATCTTCAAGAGGGCTTCTCCATCGCCAGAGGGCGTGGTGGCGGAGAA 155  
QY 21 LySThrLySGInGlyValThrGluAlaAlaGluLyThLySGInGlyValMetTyVal 40  
DB 156 AAGACCAACAGGGGGTGAAGAGACGCTGAAGAACCAAGAGGGGGTCAATGATG 215  
QY 41 G1yAlaLyThLySGInGluSnVAlValGInSerValThSerValAlaGluLyThLyS 60  
DB 216 GGAGCCAAAGACAGAGATGTTGTACAGAGCGTCACTCACTGCGCAAGAACCAAG 275  
QY 61 GluGlnAlaAsnAlaValSerIleAlaValSerValAsnThrValAlaThrLyS 80  
DB 276 GAGCAGGCCAAGCGCTGAGCGGTGTGTAGAGCGTCAACCTGTGGCCACCAAG 335  
QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValArgLySGInPleu 100  
DB 336 ACCGTGAGAGAGCGGAGAACATCGCGTCACTCCGGGGTGTGGCAAGAGACTTG 395  
QY 101 ArgProSerAlaProGInGInGluGlyGluAlaSerLySGInGluGluValAlaGlu 120  
DB 396 AGGCCATCTGCCCCCAACAGAGGGGTGAGCATCCAAAGAGAAAGAGAGTGGCAGAG 455



RESULT 7  
121 GtualaGlnSerGIyGIYaSp 127  
|=====|  
Db BQ893395

DEFINITION  
AGENCOURT 8121065 lupski\_dorsal\_root\_ganglion Homo sapiens CDNA  
clone IMAGE:6178582 5', mRNA sequence.

ACCESSION  
BQ893395

VERSION  
BQ893395.1 GI:22285409

KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)

ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 781)  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strusberg, Ph.D.  
Email: cga@bpe-remail.nih.gov  
Tissue Procurement: Dr. James R. Lupscki  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
Cloning Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.lnl.gov>  
Plate: LLM13558 row: F column: 23  
High quality sequence stop: 436.  
Location/Qualifiers  
1..781  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_image:6178582"  
/sex="male"  
/tissue\_type="dorsal root ganglia"  
/dev\_stage="adult, 36 yr"  
/lab\_host="DH10B"  
/clone\_lib="lupski\_dorsal root ganglion"  
/note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1:  
NotI; Site\_2: SalI; CDNA made by oligo-dT priming.  
Directionally cloned using the following adaptors:  
5'-TCGACCACAGCGGTCCG-3' and  
5'-ACTCATGCTTAGATCGCAGCGGCCCT(15)-3'. Size selected >  
1 kb for average insert length 1.7 kb. This is a primary  
library, non-amplified. Library constructed by Life  
Technologies and donated by J. Lupscki, M.D./Ph.D. (Baylor  
College of Medicine) and is available through Life  
Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 2,166-58 Length: 781

Score: 602.00 Matches: 125

Percent Similarity: 100.00% Conservative: 2

Best Local Similarity: 98.43% Mismatch: 0

Query Match: 98.69% Indels: 0

DB: 5 Gaps: 0

US-09-017-715A-2 (1-127) x BQ893395 (1-781)

Oy 1 MetAspValPheLySylSGlyPhseSerIleAlaLySylGlyValAlGlyAlaValGlu 20  
Db 71 ATGATGCTCTCAAGAGGGCTTCTCCATCGCCAAAGAGGGCGTGGTGGCGGTGCAA 130

Oy 21 LySThrLySgInGlyValThrGlualaAlaGluLySThrLySgInGlyValMetTyVal 40  
Db 131 AAGACCAAGCAGGGGGGTGACGGAAGCACCTGAGAAGAACCAAGAGGGGGTCATGTATGTG 190

Oy 41 GlyAlaLySThrLySgluaSnVaValGlnSerValThrSerValAlaGluLySThrLyS 60

[illegible]

US-09-017-715A-2 (1-127) x BM921124 (1-809)

QY 1 MetaspvalPheLysGlyPheSerIleAlaLysGlyValGlyAlaValGlu 20  
DB 114 ATGATGTCCTTCAAGAGGGCTTCTCCATCGCCCAAGAGGGCGTGGTGGCGGAGAA 173

QY 21 LysThrLysGlnGlyValThrGluAlaGluLysThrLysGlnGlyValMetYrVal 40  
DB 174 AAGACCAAGCAGAGGGGTGAGGAGCAGCTGAGAGAACAGAGAGGGGTCAATGATGTG 233

QY 41 GJYAlaLysThrLysGlnValValGlnSerValThrSerValAlaGluLysThrLys 60  
DB 234 GGAGCCAAACCAAGAGAGATGTTGACAGAGGTGACCTCACTGGCCGAGAAAGCAAG 293

QY 61 GJGlnAlaAsnAlaValSerLysAlaValSerSerValAsnThrValAlaThrLys 80  
DB 294 GAGCAGGCCCAAGCCCTGAGCGAGGCTGTGTGAGCAGCGTCAACACTGTGGCCACCAAG 353

QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValAlaArgLysGluAspLeu 100  
DB 354 ACCGTGAGAGAGCGAGAAACATCGCGTCACTCCGGGGTGGTGGCAAGGAGACTTG 413

QY 101 ArgProSerAlaProGlnGlnGluGluAlaSerLysGluLysGlnGluValAlaGlu 120  
DB 414 AAGCCATCTGCCCCCAACAGAGAGGTGAGGATCCAAAGAAAGAGAAAGTGGCAGAG 473

QY 121 GJAlaGlnSerGlyGlyAsp 127  
DB 474 GAGGCCCAAGTGGGGAGAGAC 494

RESULT 9  
BQ882072 855 bp mRNA linear EST 16-ANG-2002  
LOCUS AGENCOURT 8586140 lupski\_sympathetic\_trunk Homo sapiens cDNA clone  
DEFINITION IMAGE:6195522 5', mRNA sequence.  
BQ882072  
BO882072.1 GI:22274080  
EST.  
Homo sapiens (human)  
Source  
Organism Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 855)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: Dr. James R. Lupski  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM13602 row: h column: 19  
High quality sequence stop: 667.  
Location/Qualifiers  
1. 855  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6195522"  
/sex="male"  
/tissue\_type="sympathetic trunk"  
/dev stage="adult, 16 yr"  
/lab host="DH10B"  
/clone lib="lupski\_sympathetic trunk"  
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site 1:  
NotI; Site 2: SalI; cDNA made by oligo-dT priming.  
Directionally cloned using the following adaptors:  
5'-TCGACCCACGCGTCCG-3' and

FEATURES  
source

5'-GACTAGTCTAGATCGAGCGGCCCT(15)-3'. Size selected >  
1 kb for average insert length 1.9 kb. This is a primary  
library non-amplified. Library constructed by Life  
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
College of Medicine); available through Life  
Technologies."

ORIGIN

Alignment Scores:  
Pred. No.: 2,436-58 Length: 855  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.63% Indels: 0  
DB: 5 Gaps: 0

US-09-017-715A-2 (1-127) x BQ882072 (1-855)

QY 1 MetaspvalPheLysGlyPheSerIleAlaLysGlyValGlyAlaValGlu 20  
DB 94 ATGATGTCCTTCAAGAGGGCTTCTCCATCGCCCAAGAGGGGTGGTGGCGGAGAA 153

QY 21 LysThrLysGlnGlyValThrGluAlaGluLysThrLysGlnGlyValMetYrVal 40  
DB 154 AAGACCAAGCAGAGGGGTGAGGAGCAGCTGAGAGAACAGAGAGGGGTCAATGATGTG 213

QY 41 GJYAlaLysThrLysGlnValValGlnSerValThrSerValAlaGluLysThrLys 60  
DB 214 GAGCCAAAGCAGAGAGATGTTGACAGAGGTGACCTCACTGGCCGAGAAAGCAAG 273

QY 61 GJGlnAlaAsnAlaValSerLysAlaValSerSerValAsnThrValAlaThrLys 80  
DB 274 GAGCAGGCCCAAGCCCTGAGCGAGGCTGTGTGAGCAGCGTCAACACTGTGGCCACCAAG 333

QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValAlaArgLysGluAspLeu 100  
DB 334 ACCGTGAGAGAGCGAGAAACATCGCGTCACTCCGGGGTGGTGGCAAGGAGACTTG 393

QY 101 ArgProSerAlaProGlnGlnGluGluAlaSerLysGluLysGlnGluValAlaGlu 120  
DB 394 AAGCCATCTGCCCCCAACAGAGAGGTGAGGATCCAAAGAAAGAGAAAGTGGCAGAG 453

QY 121 GJAlaGlnSerGlyGlyAsp 127  
DB 454 GAGGCCCAAGTGGGGAGAGAC 474

RESULT 10  
BQ439430 884 bp mRNA linear EST 24-MAY-2002  
LOCUS AGENCOURT 7911914 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:6010114  
DEFINITION 5', mRNA sequence.  
BQ439430  
BQ439430.1 GI:21178506  
EST.  
Homo sapiens (human)  
Source  
Organism Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 884)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTF/Gazdar  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM13197 row: k column: 11  
High quality sequence stop: 559.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

# FEATURES

Location/Qualifiers  
1. .884  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6010114"  
/issue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_68"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies."

## ORIGIN

### Alignment Scores:

Pred. No.: 2,546-58 Length: 884  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 5 Gaps: 0

US-09-017-715a-2 (1-127) x BQ439430 (1-884)

Qy 1 MetaspvAlphelyslysglypHeserTleAlalyssglyValValGlyAlValGlu 20  
Db 94 ATGATGTCTTCAAGAGGGCTTCTCCATCCCAAGAGGGCGTGGTGGCGGAA 153  
Qy 21 LyethrlysglnglyValThrgluAlaGluThrylthrysglyValMetTyVal 40  
Db 154 AAGCCAGAGCGGGGCGGAGCGAGCGTGAAGAGCGAGAGGGGCTCATATGTG 213  
Qy 41 GlyAlalyThrylthrysglyValGlnSerValThrylthrysglyValGlu 60  
Db 214 GAGGCCAGAGCGAGAGGAGTGTGTACAGAGCGTGAAGCGGCGGAGAGCAAG 273  
Qy 61 GluGlnAlaAlaValSerlyAlaValSerValThrylthrysglyValGlu 80  
Db 274 GAGAGCGCCAGAGCGGGGCGGAGCGTGTGTGAGCGCGTCAACACTGTGCCAAG 333  
Qy 81 ThrValGluGluAlaGluAlaValThrylthrysglyValAlaGlyValGlu 100  
Db 334 ACCGTGAGAGAGCGGAGAGCATCGCGTCACTCCGGGGTGGTGGCGAGAGACTTG 393  
Qy 101 ArgProSerAlaProGlnGlnGlyGluAlaSerlysglyValGluValAlaGlu 120  
Db 394 AGGCCATCTGCCCGCCCAAGAGAGGTGAGGATCCCAAGAGAGAGAGTGCAGAG 453  
Qy 121 GluAlaGlnSerlyGlyAap 127  
Db 454 GAGGCCAGAGTGGGAGAG 474

## RESULT 11

LOCUS BQ221776 891 bp mRNA linear EST 02-MAY-2002  
DEFINITION ACENECOURT\_7549973 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:6057784  
5', mRNA sequence.  
BQ221776  
BQ221776.1 GI:20403176

ACCESSION BQ221776  
VERSION BQ221776  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NIH-MGC http://mgs.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

AUTHORS Contact: Robert Strausberg, Ph.D.  
TITLE Email: cga@pds-remail.nih.gov  
JOURNAL Tissue Procurement: DCTD/DRP/Gardar  
COMMENT CDNA Library Preparation: Life Technologies, Inc.

# FEATURES

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM13321 row: m column: 17  
High quality sequence stop: 734.  
Location/Qualifiers  
1. .891  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6057784"  
/issue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_68"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies."

## ORIGIN

### Alignment Scores:

Pred. No.: 2,566-58 Length: 891  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 5 Gaps: 0

US-09-017-715a-2 (1-127) x BQ221776 (1-891)

Qy 1 MetaspvAlphelyslysglypHeserTleAlalyssglyValValGlyAlValGlu 20  
Db 118 ATGATGTCTTCAAGAGGGCTTCTCCATCCCAAGAGGGCGTGGTGGCGGAA 177  
Qy 21 LyethrlysglnglyValThrgluAlaGluThrylthrysglyValMetTyVal 40  
Db 178 AAGCCAGAGCGGGGCGGAGCGTGAAGAGCGAGAGGGGCTCATATGTG 237  
Qy 41 GlyAlalyThrylthrysglyValGlnSerValThrylthrysglyValGlu 60  
Db 238 GAGGCCAGAGCGAGAGGAGTGTGTACAGAGCGTGAAGCGGCGGAGAGCAAG 297  
Qy 61 GluGlnAlaAlaValSerlyAlaValSerValThrylthrysglyValGlu 80  
Db 298 GAGAGCGCCAGAGCGGGGCGGAGCGTGTGTGAGCGTCAACACTGTGCCAAG 357  
Qy 81 ThrValGluGluAlaGluAlaValThrylthrysglyValAlaGlyValGlu 100  
Db 358 ACCGTGAGAGAGCGGAGAGCATCGCGTCACTCCGGGGTGGTGGCGAGAGACTTG 417  
Qy 101 ArgProSerAlaProGlnGlnGlyGluAlaSerlysglyValGluValAlaGlu 120  
Db 418 AGGCCATCTGCCCGCCCAAGAGAGGTGAGGATCCCAAGAGAGAGAGTGCAGAG 477  
Qy 121 GluAlaGlnSerlyGlyAap 127  
Db 478 GAGGCCAGAGTGGGAGAG 498

## RESULT 12

LOCUS BQ286466 903 bp mRNA linear EST 21-FEB-2001  
DEFINITION 602382954F1 NIH\_MGC\_93 Homo sapiens cDNA clone IMAGE:4500613 5', mRNA sequence.  
BQ286466  
BQ286466.1 GI:13039369

ACCESSION BQ286466  
VERSION BQ286466  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
(bases 1 to 903)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LHM10366 row: k column: 14  
 High quality sequence stop: 681.

FEATURES  
 source Location/Qualifiers  
 1..903  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:450613"  
 /tissue\_type="transitional cell papilloma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1lb="NIH\_MGC\_93"  
 /note="Organ: bladder; Vector: PCMV-SPORE; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.:	2,61e-58	Length:	903
Score:	602.00	Matches:	125
Percent Similarity:	100.00%	Conservative:	2
Best Local Similarity:	98.43%	Mismatches:	0
Query Match:	98.69%	Indels:	0
DB:	4	Gaps:	0

US-09-017-715A-2 (1-127) x BG286466 (1-903)

QY 1 MetAspValPheIysLysGlyPheSerIleAlaLysGlyValValGluValGlu 20  
 DB 87 ATGAGATGCTTCAAGAGGGCTTCTCCATCGCAAGAGGGCGTGGTGGCGTGA 146  
 QY 21 LysThrValGlnGlyValThrGluAlaGluValThrLysGlnGlyValMetTYrVal 40  
 DB 147 AAGACCAACAGAGGGGTGAACGAGACGCTGAGAACACAGAGAGGGGTGATGATG 206  
 QY 41 GAlAAlaLysThrLysGluValValGlnSerValThrSerValAlaGluValThrLys 60  
 DB 207 GAGGCCAAGACCAAGAGAAATGTTTACAGAGGCTGACCTGCGCCAGAAAGACCAAG 266  
 QY 61 GluGlnAlaAsnAlaValSerLysAlaValValSerSerValAsnThrValAlaThrLys 80  
 DB 267 GAGCAGGCCAACGCCCTGAGCGAGGCTGTGTGAGAGCGTCAACACTGTGGCCACCAAG 326  
 QY 81 ThrValGluGluValGluValAsnIleAlaValThrSerGlyValValAlaArgLysGluAspLeu 100  
 DB 327 ACCGTGAGAGAGCGGAGAACATCGCGGTCACTCCGGGGTGTGGCAAGAGGACTTG 386  
 QY 101 ArgProSerAlaProGlnGlnGlnGlyGluValAserLysGluValGlnGluValAlaGlu 120  
 DB 387 AGGCCATCTGCCCCCAACAGAGGGGTGAGGATCCAAAGAGAAAGAGAGTGGCAGAG 446  
 QY 121 GluAlaGlnSerGlyLysP 127  
 DB 447 GAGGCCACAGATGGGAGAGC 467  
 RESULT 13  
 B1548891 558 bp mRNA linear EST 05-SEP-2001  
 LOCUS 603189023F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5260589 5',  
 DEFINITION mRNA sequence.

ACCESSION B1548891  
 VERSION B1548891.1 GI:15436203  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: Miklos Palkevics, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LHM11656 row: m column: 06  
 High quality sequence stop: 558.

FEATURES  
 source Location/Qualifiers  
 1..558  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5260589"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /clone\_1lb="NIH\_MGC\_95"  
 /note="Organ: brain; Vector: pBluescriptPR (modified pBluescript KS+); Site\_1: BamHI, Site\_2: SalI-XhoI (96cag); Oligo-dT primed using primer 5'-TTTTTCTTTTCTTTTCTTT-3', size-selected for average insert size 2.5 kb and normalized to ROP 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NHGRI), National Institutes of Health). Note: this is a NIH\_MGC Library."

Alignment Scores:

Pred. No.:	8,82e-58	Length:	558
Score:	595.00	Matches:	124
Percent Similarity:	99.21%	Conservative:	2
Best Local Similarity:	97.64%	Mismatches:	1
Query Match:	97.54%	Indels:	0
DB:	4	Gaps:	0

US-09-017-715A-2 (1-127) x B1548891 (1-558)

QY 1 MetAspValPheIysLysGlyPheSerIleAlaLysGlyValValGluValGlu 20  
 DB 130 ATGAGATGCTTCAAGAGGGCTTCTCCATCGCAAGAGGGCGTGGTGGCGTGA 189  
 QY 21 LysThrValGlnGlyValThrGluAlaGluValThrLysGlnGlyValMetTYrVal 40  
 DB 190 AAGACCAACAGAGGGGTGAACGAGACGCTGAGAACACAGAGAGGGGTGATGATG 249  
 QY 41 GAlAAlaLysThrLysGluValValGlnSerValThrSerValAlaGluValThrLys 60  
 DB 250 GAGGCCAAGACCAAGAGAAATGTTTACAGAGCGTCACTGCTGCCAGAAAGACCAAG 309  
 QY 61 GluGlnAlaAsnAlaValSerLysAlaValValSerSerValAsnThrValAlaThrLys 80  
 DB 310 GAGCAGGCCAACGCCGTGAGCGAGGCTGTGTGAGAGCGTCAACACTGTGGCCACCAAG 369  
 QY 81 ThrValGluGluValGluValAsnIleAlaValThrSerGlyValValAlaArgLysGluAspLeu 100  
 DB 370 ACCGTGAGAGAGCGGAGAACATCGCGGTCACTCCGGGGTGTGGCAAGAGGACTTG 429

Qy	101	ArpProseralAaprogInglnglglYgluaJaseerlysglyVsglucInuValAlaActu	120
Db	430	AGGCCATGTGCCCCCAAGAGGGGTGTGTGCATCCAAAGAGAAAGAGGAAGTGCACAG	489
Qy	121	GluaIaInserGlyGLYASP	127
Db	490	GAGGCCACAGACTGGGGGAGAC	510
RESULT 14			
LOCUS	BG708703	568 bp	mRNA
DEFINITION	602674249P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4796820 5',		linear EST 07-MAY-2001
ACCESSION	BG708703		
VERSION	BG708703.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 568)		
JOURNAL	NIH-MGC <a href="http://mgs.nci.nih.gov/">http://mgs.nci.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA library Preparation: Michael J. Brownstein (NHGRI), Shiroaki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
File: LLNL0682 row: a column: 13  
High quality sequence slice: 566.

**FEATURES**  
**Source**

```

1. 568
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone_image="4196820"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_id="N1H_MGC_96"
/notes="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gcgag); Oligo-dt primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to R0.5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carnac, in
preparation). Library constructed by M. Brownstein
(NIMH/NHRI, National Institutes of Health). Note: this
is a NIH MGC library."

```

**ORIGIN**

### Alignment Scores:

```
Pred. No.: 9.02e-
Score: 995.00
Percent Similarity: 99.21%
Best Local Similarity: 97.64%
Query Match: 97.54%
DB: 4
```

US-09-017-715A-2 (1-127) X BG708703 (1-568)

QY I MetADPValPheLysGLyPheSerIleAlaLysGLyValValGlyAlaValGlu 20

DB 108 ATGATGTCTTCAAGAGGGCTTCATCGCCAAGAGGGCGTGCTGGCTGCCGTGAA 16

QY 21 LysThrLysGlnGlyValThrGluAlaAlaGluLysThrLysGlnGlyValMetTyrVal 40

Dd	168	AAGACCAACGACGAGGGGTGACCGAACCCGCTGAGAAACCAAGAGGGGGCTCATGTATGTC	227
Qy	41	GIYAlaIyThrLyEgLUbsnValValGlnSerValThrSerValAlaGluIyThrLyS	60
Dd	228	GGAGCCAAAGACCCAAAGAGAAATGTTTACAGACCGTGAACCTCAGTGGCCAGAAAGACCAAG	287
Qy	61	GIUGlnAlaAsnAlaValSerLySAlaValValSerSerValAsnThrValAlaThrLyS	80
Dd	288	GAGCAGGGCCAACGGGTGAGCGAGGCTGTGTGTGAGAGCGCTCAACACTGTGGCCACCAAG	347
Qy	81	ThrValGIUGluAlaGluAsnIleAlaValThrSerGIYValValArgLyEgLUbsnPlen	100
Dd	348	ACCGTGGAGGAGGCGGAGAAATCCCGGTCACTCCGGGGTGTGGTGCAGAGGAGACTTG	407
Qy	101	ArgProSerAlaProGlnGlnGluIyGIuAlaSerLyGIuLyEgLUgluValAlaGlu	120
Dd	408	AGGCCATCTGCCCCCAACAGGAGGGTGTGCATCCAAAGAAAGAGAAAGTGGCAGAG	467
Qy	121	GIUAlaGlnSerGIYGIYAsp	127
Dd	468	GAGGCCACAGAGTGGGGAGAC	488

RESULT 15	617 bp	RNA	linear	EST 07-MAY-2001
LOCUS	BG707764			
DEFINITION	60267110.F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:473833 5', mRNA sequence.			
ACCESSION	BG707764			
VERSION	BG707764.1	GI:13984439		

SOURCE ORGANISM	REFERENCE TITLE JOURNAL COMMENT
Homo sapiens (human)	
Homo sapiens	
Bakaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi	
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
1 (bases 1 to 617)	
NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .	
National Institutes of Health, Mammalian Gene Collection (MGC)	
Unpublished (1999)	
Contact: Robert Strausberg, Ph. D.	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 617)	NIH-MGC	<a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .	National Institutes of Health, Mammalian Gene Collection (MGC).	
		Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.	Email: <a href="mailto:cgapds-remail.nih.gov">cgapds-remail.nih.gov</a>		
		Tissue Procurement: Miklos Palkovits, M.D., Ph.D.		
		cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shioh		
		Toshiyuki and Piero Carninci (RIKEN)		
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
		DNA Sequencing by: Incyte Genomics, Inc.		
		Clone distribution: MGC clone distribution information can be		
		found through the I.M.A.G.E. Consortium/LNLN at:		
		<a href="http://image.llnl.gov">http://image.llnl.gov</a>		
	Plate: LLM010674	row: e	column: 02	
	High quality sequence stop: 617.			
FEATURES		Location/Qualifiers		
source		1..617		

```

/note="Organ: brain; Vector: pBluescriptPR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-xhoI
(gcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROP 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carinci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this
is a NIH-MGC Library."

```

**Alignment Scores:**

Pred. No.: 1e-57 Length: 617  
Score: 595.00 Matches: 124  
Percent Similarity: 99.21% Conservative: 2  
Best Local Similarity: 97.64% Mismatches: 1  
Query Match: 97.54% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715a-2 (1-127) x BG707764 (1-617)

```
QY      1 MetAEPValPheLyLeLySerIleAlaLyLeLyGlyValValGlyAlaValGlu 20
DB      58 ATGGATGCTTCAGAGAGGCTTCTCCATCGCCAGAGGGCGTGGTGGCGGTGGA 117
QY      21 LysThrLySGInGlyValThrGluAlaGluLyThrLySGInGlyValMetTyVal 40
DB      118 AAGACCAAGCAGGGGGGTGCGAGAGCACTGAGAGAACCAAGAGGGGGTCAATGTTG 177
QY      41 GIYAlaLySThrLySGInuAsnValValGlnSerValThrSerValAlaGluLySThrLy 60
DB      178 GGAGCCAAAGACCAAGAGAAATGTTGACAGAGCGTGACCTCACTGCGCGAAGAACCAAG 237
QY      61 GIuGInAlaAsnAlaValSerLySAlaValValSerSerValAsnThrValAlaThrLy 80
DB      238 GAGCAGGCGCAAGCGCGTGAAGCGCTGTGTGAGCAGCGTCAACACTGTGCGCAACCAAG 297
QY      81 ThrValGluGluValaGluAsnIleAlaValThrSerGlyValValArgLySGInuAsnLeu 100
DB      298 ACCGTGAGAGAGGCGAGAAACATCGCGGTCACTCCGGGGTGTGCGCAAGAGACTTG 357
QY      101 ArgProSerAlaProGInGInGInGlyGluAlaSerLySGInuLySGInuValAlaGlu 120
DB      358 AGGCCATCTGCCCGCCCAACAGAGGGGTGTGCATCCAAAGAGAAAGAGAGTGGCAGAG 417
QY      121 GIuAlaGlnSerGlyGly*Sp 127
DB      418 GAGGCCCAAGTGGGGGAGAC 438
```

Search completed: May 4, 2005, 13:46:32  
Job time : 5788.07 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 4, 2005, 09:07:52 ; Search time 58.4698 Seconds

(without alignments)

809.955 Million cell updates/sec

Title: US-09-017-715A-2\_COPY\_120\_127

Perfect score: 41

1 EBAQSGSD 8

## Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool.h/US9017175/runat.04052005.100743.25600/app.query.fasta\_1.661  
-DB=N.GeneSeq -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USBR=US09017175.OCGN.1.1.703.@runat.04052005.100743.25600 -NCPD=6 -ICPU=3  
-NO MMAP -LARGESOURCY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

N.GeneSeq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	41	100.0	478	3	AAF21785	AAf21785 Human bre
2	41	100.0	479	9	ACH15493	ACH15493 Human adu
3	41	100.0	488	12	ADM66887	Adm66887 Human hom
4	41	100.0	550	2	AAV42669	AAv42669 Human bre
5	41	100.0	550	3	AAA39470	AAa39470 Human HBG

6	41	100.0	550	6	ABL63343	ABl63343 Breast ca
7	41	100.0	550	6	ABV73813	ABv73813 Human gam
8	41	100.0	550	6	ABV73915	ABv73915 Human gam
9	41	100.0	550	10	AA063568	AA063568 Human amy
10	41	100.0	550	10	ADG47636	ADg47636 Human amy
11	41	100.0	720	2	AAx23997	Aax23997 Human per
12	41	100.0	720	6	AB576519	AB576519 cDNA enco
13	41	100.0	720	10	AD543864	AD543864 Human SNC
14	41	100.0	720	12	ADH54342	ADh54342 Human SNC
15	41	100.0	796	3	AAI93778	AAi93778 Human pol
16	41	100.0	796	3	AAI93784	AAi93784 Human bre
17	41	100.0	990	13	ADR98806	ADr98806 lung spec
18	41	100.0	1125	13	AD557612	AD557612 Bacterial
19	38	92.7	110000	11	ADM27081_11	Continuation (12 o
20	37	90.2	1275	4	ABL04107	ABl04107 Drosophi1
21	37	90.2	3275	4	ABL04106	ABl04106 Drosophi1
22	36	87.8	2116	10	ADB62945	ADb62945 Human CDN
23	36	87.8	3790	10	ADC39175	ADc39175 Novel hum
24	36	87.8	4062	13	AD089903	AD089903 Antagonis
25	36	87.8	4949	12	AD085477	AD085477 Human tum
26	36	87.8	4985	2	AAK00463	Aax00463 Human typ
27	36	87.8	4985	12	ADP21328	ADp21328 Gene ADcy
28	36	87.8	5054	10	ADF74204	Adf74204 Human nov
29	36	87.8	5236	10	ADC30279	ADc30279 Human nov
30	36	87.8	5372	12	AD025367	AD025367 Human sof
31	36	87.8	349980	6	AB081847	AB081847 Bifidobac
32	35	85.4	458	4	AAK76569	AAk76569 Human imm
33	35	85.4	458	4	AAK76570	AAk76570 Human imm
34	35	85.4	518	13	ACH48722	ACH48722 Cotton pr
35	35	85.4	547	12	ACH73352	ACH73352 Human gen
36	35	85.4	677	2	AAK04876	AAk04876 Human gam
37	35	85.4	918	13	ADT46476	ADt46476 Bacterial
38	35	85.4	1273	11	ACN91507	ACn91507 Breast ca
39	35	85.4	2143	10	ADC07759	ADc07759 Rice DNA
40	35	85.4	2154	10	ADC08252	ADc08252 Rice DNA
41	35	85.4	2251	4	AD160540	ADi60540 Secreted
42	35	85.4	3081	4	AB127575	AB127575 Drosophi1
43	35	85.4	4259	10	AD583369	AD583369 Toxicity
44	35	85.4	4259	10	AD582961	AD582961 Primary r
45	35	85.4	4292	10	ADE71246	ADe71246 Novel hum

## ALIGNMENTS

RESULT 1	AAF21785	standard; DNA; 478 BP.
ID	AAF21785	
XX	AAF21785;	
AC	27-MAR-2001	(first entry)
DT		
XX		
DE		Human breast and ovarian cancer associated antigen gene SEQ ID 172.
XX		
KW		Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW		neurotropic; neurprotective; antiviral; antiallergic; hepatotropic;
KW		antidiabetic; antiinflammatory; antitumor; vulnertary; anticonvulsant;
KW		antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW		Addison's disease; allergy; autoimmune haemolytic anaemia;
KW		autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW		multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW		cardiovascular disorder; wound healing; neurological disease; ds.
XX		
OS		Homo sapiens.
PN		W020005173-A1.
XX		
PD		21-SEP-2000.
XX		
PF		08-MAR-2000; 2000MO-US0005881.
XX		
PR		12-MAR-1999; 99US-0124270P.
XX		

PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM;  
XX  
XX MPI: 2000-611515/58.  
DR P-PSDB; AAB58862.  
XX  
XX New human breast and ovarian cancer associated gene sequences and the  
PT polypeptides encoded by these genes, useful in the prevention, treatment  
PT and diagnosis of cancer, immune disorders, cardiovascular disorders and  
PT neurological diseases.  
XX  
XX Claim 1; Page 609; 1299pp; English.  
XX  
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
CC associated with breast and ovarian cancer. Included in the invention are  
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
CC isolation and characterization of the DNA and protein sequences of the  
CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
CC or antagonist sequences exhibit cytostatic; immunosuppressive; neurotropic;  
CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;  
CC antiinflammatory; antitumor; vulnerary; anticonvulsant; antibacterial;  
CC antifungal; antiparasitic and cardiac activity. The polynucleotide and  
CC protein sequences are used in the diagnosis of cancer, particularly  
CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists  
CC and antagonists may also be used in the diagnosis, prevention and treatment  
CC of immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC cardiovascular disorders such as myocardial ischaemia; wound healing;  
CC neurological diseases such as cerebral anoxia and epilepsy; and  
CC infectious diseases  
XX  
SQ Sequence 478 BP; 118 A; 150 C; 113 G; 97 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 47.8 Length: 478  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0  
  
US-09-017-715A-2\_COPY\_120\_127 (1-8) x AAF21785 (1-478)  
QY 1 GIUGUAAGlnSercIyGlyasp 8  
DB 149 GAGGAGGCCCAAGATGGGGAGAC 172  
  
RESULT 2  
ACH15493  
ID ACH15493 standard; cDNA; 479 BP.  
XX  
XX ACH15493;  
XX  
XX 13-OCT-2003 (first entry)  
XX  
XX Human adult brain cDNA #2705.  
XX  
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
KW genome mapping; biodiversity; genetic disorder.  
XX  
XX Homo sapiens.  
XX  
XX US2003073623-A1.  
XX  
XX 17-APR-2003.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX

PA (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
PI MPI: 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 2705; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC segdata.uspto.gov/sequence.html?docID=20030073623  
XX  
SQ Sequence 479 BP; 120 A; 118 C; 178 G; 56 T; 0 U; 7 Other;  
  
Alignment Scores:  
Pred. No.: 47.9 Length: 479  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
  
US-09-017-715A-2\_COPY\_120\_127 (1-8) x ACH15493 (1-479)  
QY 1 GIUGUAAGlnSercIyGlyasp 8  
DB 371 GAGGAGGCCCAAGATGGGGAGAC 394  
  
RESULT 3  
ADM66887  
ID ADM66887 standard; DNA; 488 BP.  
XX  
XX ADM66887;  
XX  
XX 03-JUN-2004 (first entry)  
XX  
XX Human homologue of murine adipocyte specific gamma synuclein DNA Seq 20.  
XX  
XX human; adipocyte specific; gene; ds; adipose tissue; anti-obesity;  
KW high mobility group I-C protein; HMGI-C; obesity; leptin; ob; diabetes;  
KW adipogenesis; hypertension; cardiovascular disease; anorectic;  
KW antidiabetic; hypotensive; gamma synuclein.  
XX  
XX Homo sapiens.  
XX  
XX WO2004011618-A2.  
XX  
XX 05-FEB-2004.  
XX  
XX 29-JUL-2003; 2003WO-US023684.  
XX



XX 29-JUL-2002; 2002US-0398785P.  
PR 12-JUN-2003; 2003US-0478206P.  
XX (HMG-E-) HMG-E INC.  
XX Chada K, Chouinard R, Aehar H, Sayed AMD;  
PI WPI; 2004-143846/14.  
XX P-PSDB; ADM67167.  
XX Identifying adipocyte specific genes, useful for treating obesity or  
PT diabetes, and for identifying drug targets, by differential gene  
PT expression analysis between adipose tissue or stromal vascular tissue of  
PT mice of different genotypes.  
XX Claim 11; SEQ ID NO 20; 91pp; English.  
XX  
XX This invention relates to a novel method for identifying targets that are  
CC over-expressed in adipose tissue and as such it provides targets for anti-  
CC obesity pharmaceutical compositions. Specifically, it refers to a high  
CC mobility group 1-C protein (HMG1-C) that is associated with obesity and  
CC is epistatic to leptin, furthermore, it refers to the ob gene where an  
CC autosomal recessive trait is linked to obesity and diabetes. The present  
CC invention describes performing differential gene expression analysis  
CC between the white adipose tissue (WAT) or stromal vascular tissue (SVT)  
CC of any two different mice selected from a group consisting of wild-type,  
CC HMG1-C-/-, ob/ob, or HMG1-C-/- ob/ob genotype mice. Accordingly, using  
CC this method novel nucleotides and the encoded proteins thereof were  
CC identified that are adipocyte specific, and as such can be used for  
CC preventing adipogenesis, diagnosing and treating diabetes, obesity,  
CC hypertension and cardiovascular disease, as well as screening for  
CC compounds that can modulate or prevent adipogenesis and treat diabetes or  
CC obesity. These compositions exhibit anorectic, antidiabetic and  
CC hypotensive activities. This polynucleotide sequence is a human homologue  
XX of a murine adipocyte specific DNA sequence of the invention.  
XX  
SQ Sequence 488 BP; 127 A; 119 C; 176 G; 66 T; 0 U; 0 Other;  
XX  
XX Alignment Scores:  
Pred. No.: 48.9 Length: 488  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0  
XX  
US-09-017-715A-2\_COPY\_120\_127 (1-8) x ADM66887 (1-488)  
QY 1 GluGluAlaGlnSerGlyGlyAap 8  
Db 369 GAGGAGGCCCAAGAGTGGGGAGAC 392  
XX  
RESULT 4  
AAV42669  
ID AAV42669 standard; cDNA; 550 BP.  
XX  
AC AAV42669;  
XX  
DT 09-NOV-1998 (first entry)  
XX  
XX Human breast cancer specific gene 1 (BCSG1) cDNA.  
DE  
XX Breast cancer specific gene 1; BCSG1; human; metastasis; diagnosis;  
KW therapy; genetic marker; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH 12..395  
FT CDS /\*tag= a  
FT  
PN W09833915-A1.

XX 06-AUG-1998.  
PD  
XX 03-FEB-1998; 98MO-US001804.  
PF  
XX 03-FEB-1997; 97US-0037080P.  
PR  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Ji H, Rosen CA;  
PI WPI; 1998-446811/38.  
XX P-PSDB; AAW63123.  
DR  
XX New isolated human breast cancer specific gene - used to develop products  
PT for the diagnosis, clinical management and treatment of breast cancer and  
PT metastases.  
XX  
XX Claim 4; Fig 1; 73pp; English.  
XX  
XX This cDNA clone corresponds to the transcript of the newly identified  
CC human breast cancer specific gene 1 (BCSG1), and includes an open reading  
CC frame for a 14.2 kDa protein (see AAW63123). It was isolated from a  
CC breast cancer cDNA library following an EST search for novel genes  
CC differentially expressed in breast cancer versus healthy breast tissue.  
CC The clone is deposited at ATCC 97175 and ATCC 97856. A gradient and stage  
CC -specific BCSG1 expression has been demonstrated from virtually no  
CC detectable expression in normal or benign breast to low level and partial  
CC expression in low grade in situ breast carcinoma and high expression in  
CC infiltrating malignant breast carcinomas. BCSG1 is useful as a breast  
CC cancer progression marker. Recombinant vectors and host cells useful for  
CC recombinant production of BCSG1 polypeptides (including epitope-bearing  
CC polypeptides) are provided. BCSG1 polynucleotides, polypeptides and  
CC antibodies can be used for the detection of breast cancer cells or breast  
CC cancer metastasis, and to develop methods for the clinical management and  
XX treatment of breast cancer  
XX  
SQ Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;  
XX  
XX Alignment Scores:  
Pred. No.: 55.4 Length: 550  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0  
XX  
US-09-017-715A-2\_COPY\_120\_127 (1-8) x AAV42669 (1-550)  
QY 1 GluGluAlaGlnSerGlyGlyAap 8  
Db 369 GAGGAGGCCCAAGAGTGGGGAGAC 392  
XX  
RESULT 5  
AAA39470  
ID AAA39470 standard; DNA; 550 BP.  
XX  
AC AAA39470;  
XX  
DT 24-AUG-2000 (first entry)  
XX  
XX Human HBGBA67A DNA.  
DE  
XX Human; ADA2; cytosolic; gene therapy; treatment; cancer;  
KW amyloid-like protein; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH 12..395  
FT CDS /\*tag= a  
FT /product= "HBGBA67"  
FT  
PN

PN	US6054289-A.		
XX			
PD	25-APR-2000.		
XX			
PF	30-AUG-1996;	96US-00705771.	
XX			
PR	30-AUG-1995;	95US-0002993P.	
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Moore PA;		
XX			
DR	WPI; 2000-338491/29.		
DR	P-PSDB; AAY87779.		
XX			
PT	New polynucleotide encoding human AD2 is useful for treating cancer and for isolating cDNAs and genes having similar biological activity.		
PS			
XX	Disclosure; Col 27-28; 54pp; English.		
CC			
CC	This invention describes a novel polynucleotide (1) encoding human ADA2.		
CC	The products of the invention have cytosratic activity and can be used		
CC	for gene therapy. (1) is useful for treating cancer; as primers and		
CC	probes for isolating full length cDNA and genes having similar biological		
CC	activity. This sequence encodes a polypeptide derived from the human		
CC	H8BBA67X clone which is an amyloid-like protein found in breast tissue		
XX			
SQ	Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;		
	Alignment Scores:		
	Pred. NO.:	55.4	Length: 550
	Score:	41.00	Matches: 8
	Percent Similarity:	100.00%	Conservative: 0
	Best Local Similarity:	100.00%	Mismatches: 0
	Query Match:	100.00%	Indels: 0
	DB:	3	Gaps: 0
US-09-017-715A-2_COPY_120_127 (1-8) x AAA39470 (1-550)			
OY	1	GlutlnAlaGlnSerGlyGlyasp 8	
Db	369	GAGGAGGCCACAGAGTGCGGAGAC 392	
RESULT 6			
ABL63343			
ID	ABL63343 standard; DNA; 550 BP.		
XX			
AC	ABL63343;		
XX			
DT	15-MAY-2002 (first entry)		
XX			
DE	Breast cancer related gene sequence SEQ ID NO:1680.		
XX			
KM	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;		
KM	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;		
KM	cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;		
XX	gene; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200194629-A2.		
XX			
PD	13-DEC-2001.		
XX			
PF	30-MAY-2001; 2001WO-US010838.		
XX			
PR	05-JUN-2000; 2000US-0209473P.		
PR	05-JUN-2000; 2000US-0209531P.		
PR	18-SEP-2000; 2000US-0233133P.		
PR	18-SEP-2000; 2000US-0233617P.		
PR	20-SEP-2000; 2000US-0234009P.		
PR	20-SEP-2000; 2000US-0234034P.		
PR	20-SEP-2000; 2000US-0234052P.		
PR			

XX	PR	22-SEP-2000;	2000US-0234509P.
XX	PR	22-SEP-2000;	2000US-0234567P.
XX	PR	25-SEP-2000;	2000US-0234923P.
XX	PR	25-SEP-2000;	2000US-0234924P.
XX	PR	25-SEP-2000;	2000US-0235077P.
XX	PR	25-SEP-2000;	2000US-0235082P.
XX	PR	25-SEP-2000;	2000US-0235134P.
XX	PR	25-SEP-2000;	2000US-0235637P.
XX	PR	26-SEP-2000;	2000US-0235638P.
XX	PR	27-SEP-2000;	2000US-0235711P.
XX	PR	27-SEP-2000;	2000US-0235720P.
XX	PR	27-SEP-2000;	2000US-0235840P.
XX	PR	27-SEP-2000;	2000US-0235863P.
XX	PR	28-SEP-2000;	2000US-0236028P.
XX	PR	28-SEP-2000;	2000US-0236032P.
XX	PR	28-SEP-2000;	2000US-0236033P.
XX	PR	28-SEP-2000;	2000US-0236034P.
XX	PR	28-SEP-2000;	2000US-0236109P.
XX	PR	28-SEP-2000;	2000US-0236111P.
XX	PR	29-SEP-2000;	2000US-0236842P.
XX	PR	29-SEP-2000;	2000US-0236891P.
XX	PR	02-OCT-2000;	2000US-0237172P.
XX	PR	02-OCT-2000;	2000US-0237173P.
XX	PR	02-OCT-2000;	2000US-0237278P.
XX	PR	02-OCT-2000;	2000US-0237294P.
XX	PR	02-OCT-2000;	2000US-0237295P.
XX	PR	02-OCT-2000;	2000US-0237316P.
XX	PR	03-OCT-2000;	2000US-0237425P.
XX	PR	03-OCT-2000;	2000US-0237598P.
XX	PR	03-OCT-2000;	2000US-0237604P.
XX	PR	03-OCT-2000;	2000US-0237606P.
XX	PR	03-OCT-2000;	2000US-0237608P.
XX	PR	01-NOV-2000;	2000US-0244867P.
XX	PR	01-NOV-2000;	2000US-0245084P.
PA		(AVAL-) AVALON PHARM.	
PB		Young PB, Augustus M, Carter KC, Ebner R, Endreess G, Horrigan S,	
PI		Soppet DR, Weaver Z;	
DR		WPI; 2002-188264/24.	
XX			
PT		Screening for anti-neoplastic agent involves exposing cells to a chemical	
PT		agent to be tested for anti-neoplastic activity, and determining a change	
PT		in expression of a gene of a signature gene set.	
PS		Claim 1; SEQ ID NO 1680; 44pp; English.	
XX			
CC		The present invention describes a method (M1) for screening for an anti-	
CC		neoplastic agent. The method involves exposing cells to a chemical agent	
CC		to be tested for anti-neoplastic activity, determining a change in	
CC		expression of at least one gene (I) of a signature gene set, where (I)	
CC		comprises a sequence (S) selected from 8447 sequences (given in AB161664	
CC		to AB170110), or is at least 95% identical to (S), where a change in	
CC		expression is indicative of anti-neoplastic activity. (I) has cytosstatic	
CC		activity and can be used in gene therapy. M1 can be used for screening which is	
CC		anti-neoplastic agent, and can be used for producing a product which is	
CC		the data collected with respect to the anti-neoplastic agent as a result	
CC		of M1, and the data is sufficient to convey the chemical structure and/or	
CC		properties of the agent. M1 can be used in the treatment of cancer such	
CC		as colon, breast, stomach, lung, thyroid, ossophageal, ovarian, kidney,	
CC		prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell	
CC		cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous	
CC		cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's	
CC		tumour	
XX			
SEQ		Sequence 550 BP, 132 A, 145 C, 192 G, 81 T, 0 U, 0 Other;	
Alignment Scores:			
Pred. No.:	55.4	Length:	550
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-017-715a-2\_COPY\_120\_127 (1-8) x ABV63343 (1-550)

Qy 1 GUGUUAAGInSergIyGLYAsp 8  
Db 369 GAGGAGCCCGAGAGTGGGGAGAC 392

RESULT 7  
ABV73813  
ID ABV73813 standard; cDNA, 550 BP.

XX ABV73813;

XX 08-JAN-2003 (first entry)

XX Human gamma-synuclein Glu110 variant gene.

XX Gamma-synuclein, human; single nucleotide polymorphism; SNP;

XX schizophtenia; neuroleptic; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 12..395

XX /tag= a

XX /product= "Gamma-synuclein"

XX /transl\_except= (pos:213..215,aa:Glu)

XX replace(340,T)

XX /tag= b

XX /standard\_name= "Single nucleotide polymorphism"

XX WO200275317-A2.

XX 26-SEP-2002.

XX 14-MAR-2002; 2002WO-EP002872.

XX 15-MAR-2001; 2001US-0276306P.

XX (NOVS ) NOVARTIS AG.

XX (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX Roberts RC, Van Oostrum J, Voshol J, Tamminga CA;

XX WPI; 2002-750574/81.

XX P-PsDB; ABP54932.

XX Screening for compounds for treating or interfering with the onset of

XX Schizophrenia Spectrum Disorders, by detecting interactions of candidate

XX compounds with the gamma-synuclein polypeptide.

XX Disclosure; Fig 1; 32pp; English.

XX The present sequence is that of cDNA encoding the Glu-110 isoform of

XX human gamma-synuclein. The invention relates to an isoform of gamma-

XX synuclein that is caused by an A/T single nucleotide polymorphism (SNP)

XX at position 329 of the gamma-synuclein coding sequence. This SNP causes a

CC useful for the diagnosis of SSDs, or susceptibility to SSDs, e.g. by PCR

CC amplification of a polynucleotide encoding gamma-synuclein and analysis

CC of the occurrence of the SNP at position 329. A transgenic animal useful

CC for the study of SSDs is also claimed

XX Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 55.4 Length: 550

XX Score: 41.00 Matches: 8

XX Percent Similarity: 100.00% Conservative: 0

XX Best Local Similarity: 100.00% Mismatches: 0

XX Query Match: 100.00% Indels: 0

XX DB: 6 Gaps: 0

US-09-017-715a-2\_COPY\_120\_127 (1-8) x ABV73813 (1-550)

Qy 1 GUGUUAAGInSergIyGLYAsp 8

Db 369 GAGGAGCCCGAGAGTGGGGAGAC 392

RESULT 8  
ABV73915  
ID ABV73915 standard; cDNA, 550 BP.

XX ABV73915;

XX 08-JAN-2003 (first entry)

XX Human gamma-synuclein Val110 variant gene.

XX Gamma-synuclein, human; single nucleotide polymorphism; SNP;

XX schizophtenia; neuroleptic; mutant; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 12..395

XX /tag= a

XX /product= "Gamma-synuclein"

XX /transl\_except= (pos:213..215,aa:Glu)

XX replace(340,A)

XX /tag= b

XX /standard\_name= "Single nucleotide polymorphism"

XX WO200275317-A2.

XX 26-SEP-2002.

XX 14-MAR-2002; 2002WO-EP002872.

XX 15-MAR-2001; 2001US-0276306P.

XX (NOVS ) NOVARTIS AG.

XX (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

CC Roberts RC, Van Oostrum J, Voshol J, Tamminga CA;

XX WPI; 2002-750574/81.

XX P-PsDB; ABP54932.

XX Screening for compounds for treating or interfering with the onset of

XX Schizophrenia Spectrum Disorders, by detecting interactions of candidate

XX compounds with the gamma-synuclein polypeptide.

XX Disclosure; Page; 32pp; English.

XX The present sequence is that of cDNA encoding the Val-110 isoform of

XX human gamma-synuclein. The invention relates to an isoform of gamma-

XX synuclein that is caused by an A/T single nucleotide polymorphism (SNP)

XX at position 329 of the gamma-synuclein coding sequence. This SNP causes a

XX glutamic acid to valine change at amino acid position 110 of gamma-



CC treating genetic diseases such as muscular dystrophy or cystic fibrosis,  
CC hyperextension, angina pectoris, myocardial infarction, ulcers, asthma,  
CC allergies, psychoses, depression, migraine, vomiting, benign prostatic  
CC hypertrophy or osteoporosis. The polypeptides and polynucleotides are  
CC useful for in vitro purposes related to scientific research, synthesis of  
CC DNA and manufacture of DNA vector. The present sequence represents cDNA  
CC encoding human amyloid like protein.

XX Sequence 550 BP, 132 A, 145 C, 192 G, 81 T, 0 U, 0 Other;

#### Alignment Scores:

Pred. No.:	55.4	Length:	550
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x ADG47636 (1-550)

Qy 1 GIUG1UALAGInserGIYGIYAsp 8

Db 369 GAGGAGGCCCGAGCTGGGGGAGAC 392

#### RESULT 11

AAK29997  
ID AAK29997 standard; DNA, 720 BP.

AC AAK29997;

XX 06-JUL-1999 (first entry)

DE Human persyn gene.

XX Human; synuclein; persyn; diagnosis; neurodegenerative disorder; cancer;

XX breast; skin; intermediate filament damage; ss.

OS Homo sapiens.

PN BP908727-A1.

XX 14-APR-1999.

PF 21-SEP-1998; 98EP-00307628.

XX 19-SEP-1997; 97GB-00019879.

XX (NEUR-) NEUROPA LTD.

PA (UYSA-) UNIV ST ANDREWS.

XX WPI; 1999-217169/19.

DR P-PSDB; AAY07271.

XX New synuclein protein (persyn) and gene, useful in assays for screening,  
PT diagnosing or monitoring cancer, neurodegenerative disorders or skin  
PT disorders.

XX Claim 29; Page 16-17; 39pp; English.

XX This sequence represents the gene encoding a novel human synuclein family  
CC member designated persyn. The sequence is useful for screening,  
CC diagnosing or monitoring cancer (especially breast or skin cancer),  
CC neurodegenerative disorders or skin disorders and for identifying cells  
CC having intermediate filament damage

XX Sequence 720 BP, 173 A, 209 C, 215 G, 123 T, 0 U, 0 Other;

#### Alignment Scores:

Pred. No.:	73.5	Length:	720
Score:	41.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

DB: 2 Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x AAK29997 (1-720)

Qy 1 GIUG1UALAGInserGIYGIYAsp 8

Db 406 GAGGAGGCCCGAGCTGGGGGAGAC 429

#### RESULT 12

ABK76519  
ID ABK76519 standard; cDNA, 720 BP.

AC ABK76519;

DT 11-DEC-2002 (first entry)

DE cDNA encoding human ovarian cancer marker OV60.

XX Human; ovarian cancer; marker; cancer; familial history; brain disorder;

XX central nervous system disorder; bacterial meningitis; viral meningitis;

XX Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;

XX brain herniation; inflammation; encephalitis; testicular disorder;

XX nontuberculous granulomatous orchitis; connective tissue disorder;

XX heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;

XX histological type; carcinogenic; ovarian cancer marker; gene; ss.

XX Homo sapiens.

OS WO200271928-A2.

PN 19-SEP-2002.

PD 14-MAR-2002; 2002MO-US007826.

XX 14-MAR-2001; 2001US-0276025P.

XX 14-MAR-2001; 2001US-0276026P.

PR 10-AUG-2001; 2001US-0311732P.

PR 19-SEP-2001; 2001US-0323580P.

PR 26-SEP-2001; 2001US-0324967P.

PR 26-SEP-2001; 2001US-0325102P.

XX 26-SEP-2001; 2001US-0325149P.

XX (MIL-) MILENITUM PHARM INC.

PA Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;

PI Meyer RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;

PI Baet RC, Lu K, Schmandt RB, Zhao X, Glat K;

XX WPI; 2002-723277/78.

DR P-PSDB; ABG96420.

XX Assessing whether a patient is afflicted with ovarian cancer, useful in  
PT assessing the stage or progression of the disease, compares comparing  
PT the expression level of a cancer marker in a sample from a patient and  
PT from a non cancer patient.

XX Disclosure; Page 411; 481pp; English.

XX The present invention relates to a new method for assessing whether a  
CC patient is afflicted with ovarian cancer. The method involves comparing  
CC the expression level of a marker in a patient sample and the normal level  
CC of expression of the marker in a control non-ovarian cancer sample, where  
CC the marker is selected from 363 cancer markers described in the  
CC specification. The method of the invention is useful in diagnosing or  
CC characterizing cancer, in detecting the presence of cancer as early as  
CC possible, and the recurrence of ovarian cancer. The method may also be of  
CC particular use with patients having an enhanced risk of developing  
CC ovarian cancer (e.g. patients having a familial history of ovarian  
CC cancer). The cancer markers may be used in the management and treatment  
CC of e.g. brain and central nervous system disorders (e.g. bacterial and  
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain  
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),  
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),

CC testicular disorders (e.g. nonbuberculous granulomatous orchitis),  
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart  
CC disease or atherosclerosis). The compositions and methods may also be  
CC used in assessing the histological type of neoplasm associated with  
CC ovarian cancer, monitoring the progression of ovarian cancer, determining  
CC whether ovarian cancer has metastasized or is likely to metastasize,  
CC selecting a composition for inhibiting ovarian cancer, assessing the  
CC ovarian carcinogenic potential of a compound, or inhibiting ovarian  
CC cancer or at risk of developing ovarian cancer. The present nucleic acid  
CC sequence encodes one of the ovarian cancer markers described in the  
CC invention

CC Sequence 720 BP; 173 A; 209 C; 215 G; 123 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 73.5 Length: 720  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x ABS76519 (1-720)

OY 1 GIUGUAGAGInserGlyGlyASP 8

DB 406 GAGGAGGCCAGAGTGGGGAGAC 429

RESULT 13  
ADE43864  
ID ADE43864 standard; cDNA; 720 BP.

AC ADE43864;  
XX  
XX 29-JAN-2004 (first entry)

DE Human SNCG cDNA, SEQ ID 469.

XX Neurodegenerative disease; uPA; SNCG; IDE; KNSL1; LIPA; TNFRSF6;  
KM Alzheimer's disease; neuroprotective; neurotropic; gene therapy;  
XX Chromosome 10; gene; ss.

OS Homo sapiens.  
XX  
XX MO2003054143-A2.

PD 03-JUL-2003.

PF 25-OCT-2002; 2002MO-US034679.

XX 25-OCT-2001; 2001US-0339525P.  
PR 08-NOV-2001; 2001US-0336829P.  
PR 08-NOV-2001; 2001US-0338010P.  
PR 09-NOV-2001; 2001US-0338633P.  
PR 04-DEC-2001; 2001US-0337052P.  
PR 28-MAR-2002; 2002US-0368919P.

XX (NEUR-) NEUROGENETICS INC.  
PA (GEHO) GEN HOSPITAL CORP.

PI Becker KD, Velicelbi G, Elliott KJ, Wang X, Tanzi RE, Bertram L;  
PI Saunders AJ, Mullin KM, Sampson AJ, Blacker DL;

XX WPI; 2003-559131/52.

XX Determining a predisposition for or the occurrence of neurodegenerative  
PT disease, e.g. Alzheimer's disease by detecting in a target nucleic acid  
PT the presence or absence of an allelic variant of one or more polymorphic  
PT regions.

PS Claim 84; Page 740; 848pp; English.

XX The present invention relates to a method (M1) for determining a

CC predisposition for or the occurrence of neurodegenerative disease in a  
CC subject. The method comprises detecting in a target nucleic acid obtained  
CC from the subject the presence or absence of an allelic variant of one or  
CC more polymorphic regions of one or more genes selected from uPA  
CC (urokinase plasminogen activator), SNCG (gamma-synuclein), IDE (insulin-  
CC degrading enzyme), KNSL1 (kinesin-like protein 1), LIPA (lysosomal acid  
CC lipase), and TNFRSF6 (Tumour Necrosis Factor Receptor-SF6), where the  
CC presence of at least one of the allelic variant of one or more  
CC polymorphic regions is indicative of a predisposition for or the  
CC occurrence of neurodegenerative disease. The genes are all located on  
CC chromosome 10. M1 is useful for determining a predisposition for or the  
CC occurrence of, and for treating neurodegenerative disease, particularly  
CC Alzheimer's disease.

XX Sequence 720 BP; 172 A; 205 C; 212 G; 120 T; 0 U; 11 Other;

Alignment Scores:  
Pred. No.: 73.5 Length: 720  
Score: 41.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-09-017-715A-2\_COPY\_120\_127 (1-8) x ADE43864 (1-720)

OY 1 GIUGUAGAGInserGlyGlyASP 8

DB 406 GAGGAGGCCAGAGTGGGGAGAC 429

RESULT 14  
ADH54342  
ID ADH54342 standard; cDNA; 720 BP.

AC ADH54342;  
XX  
XX 25-MAR-2004 (first entry)

DE Human SNCG gene cDNA sequence SegID469.

XX human; neurodegenerative disease; urokinase plasminogen activator; uPA;  
KM gamma-synuclein; SNCG; insulin degrading enzyme; IDE;  
XX kinesin-like protein 1; KNSL1; lysosomal acid lipase; LIPA;  
KM tumour necrosis factor receptor SF6; TNFRSF6; Alzheimer's disease; ss.

OS Homo sapiens.

XX US2003224380-A1.

PD 04-DEC-2003.

PF 25-OCT-2002; 2002US-00282174.

XX 25-OCT-2001; 2001US-0339525P.  
PR 02-NOV-2001; 2001US-0348065P.  
PR 02-NOV-2001; 2001US-0336833P.  
PR 08-NOV-2001; 2001US-0336829P.  
PR 08-NOV-2001; 2001US-0338010P.  
PR 09-NOV-2001; 2001US-0338633P.  
PR 04-DEC-2001; 2001US-0337052P.  
PR 28-MAR-2002; 2002US-0368919P.

XX (GEHO) GEN HOSPITAL CORP.

PI Becker KD, Velicelbi G, Elliott KJ, Wang X, Tanzi RE;  
PI Bertram L, Saunders AJ, Mullin KM, Sampson AJ;

XX WPI; 2004-060538/06.

XX Determining a predisposition for or the occurrence of neurodegenerative  
PT disease, particularly Alzheimer's disease, comprises determining the  
PT presence of a polymorphism in the uPA, SNCG, IDE, KNSL1, LIPA or TNFRSF6  
PT gene.

XX Claim 84; SEQ ID NO 463; 205bp; English.  
 XX  
 CC This invention relates to a novel method of determining a predisposition  
 CC for or the occurrence of neurodegenerative disease comprising detecting  
 CC in a target nucleic acid obtained from the subject the presence of an  
 CC allelic variant of polymorphic regions of human genes selected from  
 CC uridine kinase plasmidogen activator (uPA), gamma-globulin (SNC), insulin  
 CC degrading enzyme (IDE), kinesin-like protein 1 (KSL1), lysosomal acid  
 CC lipase (LIPA) and tumour necrosis factor receptor SF6 (TNFRSF6). The  
 CC method is useful in determining the presence or predisposition to a  
 CC neurodegenerative disease, particularly Alzheimer's disease. The present  
 CC sequence is the cDNA sequence of the human SNCG gene which is related to  
 CC the invention.  
 XX  
 SQ Sequence 720 BP; 172 A; 205 C; 212 G; 120 T; 0 U; 11 Other;  
 Alignment Scores:  
 Pred. No.: 73.5 Length: 720  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-017-715A-2\_COPY\_120\_127 (1-8) x ADHS4342 (1-720)  
 QY 1 GluGluAlaGlnSerGlyGlyAsp 8  
 DB 406 GAGGAGGCCCGAGAGTGGGGGAGAC 429  
 RESULT 15  
 AAI93778  
 ID AAI93778 standard; cDNA; 783 BP.  
 XX AAI93778;  
 AC  
 XX  
 DT 06-NOV-2001 (first entry)  
 DE Human polynucleotide SEQ ID NO 13838.  
 XX  
 KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KM tissue growth factor; immunomodulatory; cancer; leukemia;  
 KM nervous system disorders; arthritis; inflammation; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164835-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 26-FEB-2001; 2001WO-US004927.  
 XX  
 PR 28-FEB-2000; 2000US-00515126.  
 PR 18-MAY-2000; 2000US-00577409.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 DR WPI: 2001-514838/56.  
 DR P-PSDB; AAO13847.  
 XX  
 PT isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 XX and treating e.g. leukemia, inflammation and immune disorders.  
 PS Claim 1; SEQ ID NO 13838; 1399bp + Sequence Listing; English.  
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO00010-AAO013910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pcl\_sequences  
 XX  
 SQ Sequence 783 BP; 187 A; 232 C; 237 G; 127 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 80.3 Length: 783  
 Score: 41.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0  
 US-09-017-715A-2\_COPY\_120\_127 (1-8) x AAI93778 (1-783)  
 QY 1 GluGluAlaGlnSerGlyGlyAsp 8  
 DB 471 GAGGAGGCCCGAGAGTGGGGGAGAC 494  
 Search completed: May 4, 2005, 09:26:22  
 Job time : 61.4638 secs

**THIS PAGE LEFT BLANK**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 4, 2005, 09:07:55 ; Search time 294.06 Seconds  
(without alignments)  
706.682 Million cell updates/sec

Title: US-09-017-715A-2

Perfect score: 610  
Sequence: 1 MDVFKKGFSIAKKGVGAVE.....EGEAKKEKVEAERAGSGD 127

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h  
-Q=/cg2\_1/USPTO.spool.h/US9017715/runat.04052005.100745.25632/app\_query.fasta\_1.661  
-DB=Issued\_Patents\_NA -OPMT=fastap -SUFFIX=rnt -MINMATCH=0.1 -LOOPT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US9017715.0CGN\_1\_116.0runat.04052005.100745.25632 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Issued\_Patents\_NA.\*
- 2: /cg2\_6/ptodata/1/ina/5A.COMB.seq.\*
- 3: /cg2\_6/ptodata/1/ina/5B.COMB.seq.\*
- 4: /cg2\_6/ptodata/1/ina/6A.COMB.seq.\*
- 5: /cg2\_6/ptodata/1/ina/6B.COMB.seq.\*
- 6: /cg2\_6/ptodata/1/ina/6C.COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	610	100.0	550	3	US-08-705-771-1
2	610	100.0	550	4	US-09-417-540-1
3	602	98.7	702	4	US-09-949-016-1915
4	595	97.5	720	4	US-09-949-016-442
5	468.5	76.8	786	5	PCT-US95-08295-1
6	307.5	50.4	1543	4	US-09-949-016-58
7	307.5	50.4	1543	4	US-09-949-016-3084
8	307.5	50.4	1543	4	US-09-949-016-3085
9	307.5	50.4	1560	5	PCT-US94-09789-1
10	303.5	49.8	1096	4	US-09-949-016-3086
11	303.5	49.8	1096	4	US-09-949-016-3087
12	284.5	46.6	703	4	US-09-949-016-2926

13	282	46.2	313	4	US-09-513-999C-2232	Sequence 2232, Ap
14	201.5	33.0	8607	4	US-09-949-016-13657	Sequence 13657, A
15	201.5	33.0	8608	4	US-09-949-016-12184	Sequence 12184, A
16	187	30.7	419	4	US-09-621-976-609	Sequence 609, App
17	184	30.2	409	4	US-09-513-999C-2347	Sequence 2347, Ap
18	171	28.0	8281	4	US-09-949-016-14668	Sequence 14668, A
19	155	25.4	113876	4	US-09-949-016-14828	Sequence 14828, A
20	155	25.4	113876	4	US-09-949-016-14829	Sequence 14829, A
21	155	25.4	115508	4	US-09-949-016-11800	Sequence 11800, A
22	155	25.4	115508	4	US-09-949-016-14826	Sequence 14826, A
23	155	25.4	115508	4	US-09-949-016-18715	Sequence 18715, A
24	113	18.5	601	4	US-09-949-016-111590	Sequence 111590, A
25	113	18.5	601	4	US-09-949-016-111738	Sequence 111738, A
26	113	18.5	601	4	US-09-949-016-111884	Sequence 111884, A
27	113	18.5	601	4	US-09-949-016-111884	Sequence 111884, A
28	113	18.5	601	4	US-09-949-016-111884	Sequence 111884, A
29	98	16.1	845	5	PCT-US96-05320A-1342	Sequence 1342, Ap
30	98	16.1	845	5	PCT-US96-05320A-1342	Sequence 25, Appl
31	98	16.1	845	3	US-08-743-6378-25	Sequence 25, Appl
32	98	16.1	1830121	4	US-09-557-884-1	Sequence 1, Appl
33	98	16.1	1830121	4	US-09-557-884-1	Sequence 1, Appl
34	98	16.1	1830121	4	US-09-643-990A-1	Sequence 1, Appl
35	98	16.1	1830121	4	US-09-643-990A-1	Sequence 1, Appl
36	95	15.6	729	4	US-09-543-681A-1761	Sequence 1761, Ap
37	94	15.4	1695	2	US-08-216-894-1	Sequence 1, Appl
38	94	15.4	1695	2	US-09-115-746-1	Sequence 7, Appl
39	94	15.4	1932	2	US-08-216-894-7	Sequence 7, Appl
40	94	15.4	1932	3	US-09-115-746-7	Sequence 2, Appl
41	94	15.4	5361	3	US-08-973-462-2	Sequence 1, Appl
42	94	15.4	6152	1	US-08-973-462-1	Sequence 1903, Ap
43	93	15.2	6717	4	US-09-107-433-1903	Sequence 1385, Ap
44	92.5	15.2	2787	4	US-09-134-000C-3185	Sequence 3423, Ap
45	92	15.1	3506	4	US-09-710-279-3423	

ALIGNMENTS

RESULT 1  
US-08-705-771-1  
Sequence 1, Application US/08705771  
Patent No. 6054289  
GENERAL INFORMATION:  
APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,  
TITLE OF INVENTION: Human Genes, Sequences and  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/705,771  
FILING DATE: August 30, 1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: MULINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-346 (PFI96)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 973-994-1744  
TELEFAX: 973-994-1744  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 550 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-705-771-1

## Alignment Scores:

Pred. No.:	2,15e-69	Length:	550
Score:	610.00	Matches:	127
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
	3	Gaps:	0

US-09-017-715A-2 (1-127) x US-08-705-771-1 (1-550)

QY 1 Metaspvalphelylysglypheserillealalylysglyvalvalglvalglu 20  
DB 12 ATGAGATGTTTCAAGAGGGCTTCTCCATCGCCAAAGAGGGGTGGTGGTGGAA 71  
QY 21 LysThrlyGlnGlyValThrGluAlaGluLysThrlyGlnGlyValMetYrVal 40  
DB 72 AAGACCAAGCAGGGGTGACGGAAGCAGCTGAGAACCAAGAGGGGTGATGTGTG 131  
QY 41 G1yAlaLysThrlyGlnGluAsnValGlnSerValThrSerValAlaGluLysThrlys 60  
DB 132 GAGGCCAAGACCAAGAGAGATGTTGTACAGAGCTGACTGAGCCCAAGAACCAAG 191  
QY 61 G1uGlnAlaAsnAlaValSerlyAlaValSerSerValAsnThrValAlaThrlys 80  
DB 192 GAGCAGGCCAAGCCGCTGAGCAAGCTGTGTGACAGCTCAACACTGTGGCCACCAAG 251  
QY 81 ThrValGlnGluAlaGluAsnIleAlaValThrSerGlyValValArglyGluAspLeu 100  
DB 252 ACCGTGAGAGAGCGAGAACATCGCGGTCACTCGGGGTGTGGCAAGAGAGACTTG 311  
QY 101 AcgProSerAlaProGlnGlnGlnGlyGluAlaSerlyGluLysGlnGluValAlaGlu 120  
DB 312 AGGCCATCTGCCCCCAACAGAGAGGTGAGCATCCAAAGAAAGAGAAAGTGGCAGAG 371  
QY 121 G1uAlaGlnSerGlyGlyAsp 127  
DB 372 GAGGCCAAGATGGGGAGAGAC 392

## RESULT 2

US-09-417-540-1  
Sequence 1, Application US/09417540  
Patent No. 6639052

## GENERAL INFORMATION:

APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,  
Jian Ni and Jing-Shan Hu  
TITLE OF INVENTION: Human Genes, Sequences and  
Expression Products  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/417,540  
FILING DATE: 14-Oct-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/705,771  
FILING DATE: August 30, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-346 (PFI96)

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 973-994-1700  
TELEFAX: 973-994-1744

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 550 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: DNA

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-417-540-1

Alignment Scores:			
Pred. No.:	2,15e-69	Length:	550
Score:	610.00	Matches:	127
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
	4	Gaps:	0

US-09-017-715A-2 (1-127) x US-09-417-540-1 (1-550)

QY 1 Metaspvalphelylysglypheserillealalylysglyvalvalglvalglu 20  
DB 12 ATGAGATGTTTCAAGAGGGCTTCTCCATCGCCAAAGAGGGGTGGTGGTGGAA 71  
QY 21 LysThrlyGlnGlyValThrGluAlaGluLysThrlyGlnGlyValMetYrVal 40  
DB 72 AAGACCAAGCAGGGGTGACGGAAGCAGCTGAGAACCAAGAGGGGTGATGTGTG 131  
QY 41 G1yAlaLysThrlyGlnGluAsnValGlnSerValThrSerValAlaGluLysThrlys 60  
DB 132 GAGGCCAAGACCAAGAGAGATGTTGTACAGAGCTGACTGAGCCCAAGAACCAAG 191  
QY 61 G1uGlnAlaAsnAlaValSerlyAlaValSerSerValAsnThrValAlaThrlys 80  
DB 192 GAGCAGGCCAAGCCGCTGAGCAAGCTGTGTGACAGCTCAACACTGTGGCCACCAAG 251  
QY 81 ThrValGlnGluAlaGluAsnIleAlaValThrSerGlyValValArglyGluAspLeu 100  
DB 252 ACCGTGAGAGAGCGAGAACATCGCGGTCACTCGGGGTGTGGCAAGAGAGACTTG 311  
QY 101 AcgProSerAlaProGlnGlnGlnGlyGluAlaSerlyGluLysGlnGluValAlaGlu 120  
DB 312 AGGCCATCTGCCCCCAACAGAGAGGTGAGCATCCAAAGAAAGAGAAAGTGGCAGAG 371  
QY 121 G1uAlaGlnSerGlyGlyAsp 127  
DB 372 GAGGCCAAGATGGGGAGAGAC 392

## RESULT 3

US-09-949-016-1915  
Sequence 1915, Application US/09949016  
Patent No. 6812339

## GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
CURRENT APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 1915  
 ; LENGTH: 702  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-1915

Alignment Scores:  
 Pred. No.: 3,29e-68 Length: 702  
 Score: 602.00 Matches: 125  
 Percent Similarity: 100.00% Conservative: 2  
 Best Local Similarity: 98.43% Mismatches: 0  
 Query Match: 98.69% Indels: 0  
 DB: 4 Gaps: 0

US-09-017-715A-2 (1-127) x US-09-949-016-1915 (1-702)

QY 1 MetAspValPheIysGlyPheSerIleAlaIysGlyValAlaIgluValAlaIglu 20  
 DB 48 ATGATGCTTCAAGAGGCTTCTCCATCGCCAAAGAGGCGGTGGTGCGTGA 107  
 QY 21 LysThrIysGlnGlyValIThrGluAlaIgluIysThrIysGlnGlyValMetTyVal 40  
 DB 108 AAGCCAAAGCAGGGGCGTGAAGAGCAGCTGAGAAAGCAAGAGGGGTCTATGTG 167  
 QY 41 GAlaIalysThrIysGluAenValIgluSerValIThrSerValAlaIgluIysThrLys 60  
 DB 168 GAGGCCAAGCAGGGAATGTGTACAGGCTGACCTCAGCGCCGAGAAAGCAAG 227  
 QY 61 GAluGlnAlaAenValIserIysAlaIValIserSerValAenThrValAlaIThrLys 80  
 DB 228 GAGCAGGCCAAGCGCGTGAAGGCTGTGTGAGCAGCGCAACCTGTCAGCCAG 287  
 QY 81 ThrValIgluIuIaIgluAenIleAlaIValIThrSerIysValAlaIgluIysGluAen 100  
 DB 288 ACCGTGAGAGAGGCGGAGAACATCGCGTCACTCCGGGGTGGTCCCAAGAGACTTG 347  
 QY 101 ArgProSerAlaProGlnGlnGlyGluAlaSerIysGluIysGlnGluAlaIglu 120  
 DB 348 AGGCCATCTCCCCCACAAGAGGCTGAGGATCCAAAGAAAGAGAAAGTGGCAGAG 407  
 QY 121 GluAlaGlnSerIysGlyAap 127  
 DB 408 GAGGCCAAGTGGGGAGAC 428

# RESULT 4

; Sequence 442, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: C1001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 442  
 ; LENGTH: 720  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-442

Alignment Scores:

Pred. No.: 2,72e-67 Length: 720  
 Score: 595.00 Matches: 124  
 Percent Similarity: 99.21% Conservative: 2  
 Best Local Similarity: 97.64% Mismatches: 1  
 Query Match: 97.54% Indels: 0  
 DB: 4 Gaps: 0

US-09-017-715A-2 (1-127) x US-09-949-016-442 (1-720)

QY 1 MetAspValPheIysGlyPheSerIleAlaIysGlyValAlaIgluValAlaIglu 20  
 DB 49 ATGATGCTTCAAGAGGCTTCTCCATCGCCAAAGAGGCGGTGGTGCGTGA 108  
 QY 21 LysThrIysGlnGlyValIThrGluAlaIgluIysThrIysGlnGlyValMetTyVal 40  
 DB 109 AAGCCAAAGCAGGGGCGTGAAGAGCAGCTGAGAAAGCAAGAGGGGTCTATGTG 168  
 QY 41 GAlaIalysThrIysGluAenValIgluSerValIThrSerValAlaIgluIysThrLys 60  
 DB 169 GAGCCAAAGCAGGGAATGTGTACAGGCTGACCTCAGTGGCCGAGAAAGCAAG 228  
 QY 61 GAluGlnAlaAenValIserIysAlaIValIserSerValAenThrValAlaIThrLys 80  
 DB 229 GAGCAGGCCAAGCGGTGAAGAGCAGCTGTGTGAGCAGCGTCAACCTGTCAGCCAG 288  
 QY 81 ThrValIgluIuIaIgluAenIleAlaIValIThrSerIysValAlaIgluIysGluAen 100  
 DB 289 ACCGTGAGAGAGGCGGAGAACATCGCGTCACTCCGGGGTGGTCCCAAGAGACTTG 348  
 QY 101 ArgProSerAlaProGlnGlnGlyGluAlaSerIysGluIysGlnGluAlaIglu 120  
 DB 349 AGGCCATCTCCCCCACAAGAGGCTGAGGATCCAAAGAAAGAGAAAGTGGCAGAG 408  
 QY 121 GluAlaGlnSerIysGlyAap 127  
 DB 409 GAGGCCAAGTGGGGAGAC 429

# RESULT 5

; Sequence 1, Application PC/TUS9508295  
 ; GENERAL INFORMATION:  
 ; APPLICANT:  
 ; TITLE OF INVENTION: BREAST SPECIFIC GENES AND PROTEINS  
 ; NUMBER OF SEQUENCES: 30  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: IBM PC compatible  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/08295  
 ; FILING DATE: 30-JUN-1995  
 ; CLASSIFICATION:  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 786 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 PCT-US95-08295-1

Alignment Scores:  
 Pred. No.: 5,95e-51 Length: 786  
 Score: 468.50 Matches: 111  
 Percent Similarity: 86.15% Conservative: 1  
 Best Local Similarity: 85.38% Mismatches: 15  
 Query Match: 76.80% Indels: 4  
 DB: 5 Gaps: 1

US-09-017-715A-2 (1-127) x PCT-US95-08295-1 (1-786)

QY 1 MetAspValPheIysGlyPheSerIleAlaIysGlyValAlaIgluValAlaIglu 20

```
DB 95 ATGGATGTTTCAGAGGCGCTTCATCGCCAGAGGCGGTGGTGGCGAGAA 154
QY 21 LYSTHLYSGINGLYVALTHRGUAAAGLULYSTHLYSGINGLYVALMETYRVAL 40
DB 155 AAGACCAACAGAGGGGTGACGAGAGCTGAGAAACCAAGAGGGGTCAATGATGTG 214
QY 41 GYALALYSTHLYSGUASN--VAL-VALGINSERVALTHRSERVALAGLULYSTH 59
DB 215 GGAGGCCAACAAGAGAGATGTTGATATGACGACCGTCACTGTCGGCGAGAGAC 274
QY 59 RLYSGUGUAAALASNALVALSERLYSALVALVALSERSERVALASRTHVALALATH 79
DB 275 CAAGGCGAGCGCCAGACGCGGTGAGCAAGCTGTGTGACAGCGTCACACTKTGGCCAC 334
QY 79 RLYSTHVALIGUGUAAAGLULASN1LEALVALTHRSERGLYVALARGLYSGUAS 99
DB 335 CAAGACCGTGAAGAGAGCGGAGAACATCGCGGTCAACCTCGGGGTGCTGCGCAAGAGGA 394
QY 99 PLEUATGPROSERVALAPROGLINGLUGLYVALASERLYSGLU-LYSGUGUVALA 119
DB 395 YTTKAGCCCATY-TKCCCCCAACAGAGGGGTGAGCATMAAAGANARAGAKXGSAAGWG 453
QY 119 IAGLUGUAAAGLINSERGLYGLYASP 127
DB 454 CWRAKKRGMSCAAGGTGGGGGAGAC 479
```

## RESULT 6

```
US-09-949-016-58
/ Sequence 58, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 58
/ LENGTH: 1543
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-58
```

## Alignment Scores:

Pred. No.:	8.5e-30	Length:	1543
Score:	307.50	Matches:	68
Percent Similarity:	70.91%	Conservative:	10
Best Local Similarity:	61.82%	Mismatches:	29
Query Match:	50.41%	Indels:	3
DB:	4	Gaps:	1

US-09-017-715A-2 (1-127) x US-09-949-016-58 (1-1543)

```
QY 1 MetaspvalPheLysGlyPheSerIleAluLysGlyValValGlyValAlu 20
DB 47 ATGGATGTTTCATGAAAGCACTTTAAAGGCCAAGAGAGGTGTGCTGCTGTGAG 106
QY 21 LYSTHLYSGINGLYVALTHRGUAAAGLULYSTHLYSGINGLYVALMETYRVAL 40
DB 107 AAAACCAACAGAGGTGTGACAGAGCAGCAAGAAACAAAGAGGTCTCTATGTA 166
QY 41 GYALALYSTHLYSGUASNVALGINSERVALTHRSERVALAGLULYSTHLYS 60
DB 167 GGCTCCAAAACAGAGAGAGTGTGTCATGTGTGCAACAGTGTGAGAAACCAA 226
```

```
QY 61 GIUGUAAALASNALVALSERLYSALVALVALSERSERVALASRTHVALALATHRYS 80
DB 227 GAGCAAGAGCAAAAGTTTGAGAGACAGTGTGAGAGGGGTGTACAGCACTAGCCAGAA 286
QY 81 THVALIGUGUAAAGLULASN1LEALVALTHRSERGLYVALARGLYSGULASPLEU 100
DB 287 ACAGTGAAGGAGAGGAGGACATTTGCACAGCCACTGCTTGTCTCAAAAAGACCAAGTTG 346
QY 101 -----ArgProSerAlaProGlnGln 107
DB 347 GGCAAGAAATGAAGAGAGCCCAAGGAA 376
```

## RESULT 7

```
US-09-949-016-3084
/ Sequence 3084, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3084
/ LENGTH: 1543
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-3084
```

## Alignment Scores:

Pred. No.:	8.5e-30	Length:	1543
Score:	307.50	Matches:	68
Percent Similarity:	70.91%	Conservative:	10
Best Local Similarity:	61.82%	Mismatches:	29
Query Match:	50.41%	Indels:	3
DB:	4	Gaps:	1

US-09-017-715A-2 (1-127) x US-09-949-016-3084 (1-1543)

```
QY 1 MetaspvalPheLysGlyPheSerIleAluLysGlyValValGlyValAlu 20
DB 47 ATGGATGTTTCATGAAAGCACTTTCAAGGCCAAGAGAGGTGTGCTGCTGCTGTGAG 106
QY 21 LYSTHLYSGINGLYVALTHRGUAAAGLULYSTHLYSGINGLYVALMETYRVAL 40
DB 107 AAAACCAACAGAGGTGTGACAGAGCAGCAAGAAACAAAGAGGTCTCTATGTA 166
QY 41 GYALALYSTHLYSGUASNVALGINSERVALTHRSERVALAGLULYSTHLYS 60
DB 167 GGCTCCAAAACAGAGAGAGTGTGTCATGTGTGCAACAGTGTGAGAAACCAA 226
QY 61 GIUGUAAALASNALVALSERLYSALVALVALSERSERVALASRTHVALALATHRYS 80
DB 227 GAGCAAGAGCAAAAGTTTGAGAGACAGTGTGAGAGGGGTGTACAGCACTAGCCAGAA 286
QY 81 THVALIGUGUAAAGLULASN1LEALVALTHRSERGLYVALARGLYSGULASPLEU 100
DB 287 ACAGTGAAGGAGAGGAGGACATTTGCACAGCCACTGCTTGTCTCAAAAAGACCAAGTTG 346
QY 101 -----ArgProSerAlaProGlnGln 107
DB 347 GGCAAGAAATGAAGAGAGCCCAAGGAA 376
```

## RESULT 8

```

US-09-949-016-3085
; Sequence 3085, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3085
; LENGTH: 1543
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3085

Alignment Scores:
Pred. No.: 8.5e-30 Length: 1543
Score: 307.50 Matches: 68
Percent Similarity: 70.91% Conservative: 10
Best Local Similarity: 61.82% Mismatches: 29
Query Match: 50.41% Indels: 3
DB: 4 Gaps: 1

US-09-017-715A-2 (1-127) x US-09-949-016-3085 (1-1543)
QY 1 MetAspValPheIySgIyPheSerIleAlaIySgIyValIglIyAlaIglu 20
DB 47 ATGGATGATATTCAAGAGCACTTCAAGGCCAAGGAGGAGTGTGGCTGCTGAG 106
QY 21 LyeThrIySgIyValIthrGluAlaIgluIyThrIySgIyValIgluIyMetTyVal 40
DB 107 AAAACCAACAGGCTGTGGCAGAGCAGGAAAGCAAAAGGCTGTCTATGTA 166
QY 41 G1AlaIyThrIySgIyValIthrGluAlaIgluIyThrIySgIyValIgluIyThrIyS 60
DB 167 GGCTCCAAACCAAGGAGGAGGAGTGTGATGCTGTGGCAACAGTGGCTGAGAAACCAA 226
QY 61 GluGlnAlaAsnAlaValSerIyAlaValIySerSerValIAsnThrValAlaThrIyS 80
DB 227 GAGCAAGTGAACAATGTTGAGAGCAGTGTGACGGGTGACAGCAGTAGCCAGAG 286
QY 81 ThyValIgluIyAlaIgluAsnIleAlaValIthrSerIyValIyAlaIygluIyAspIleu 100
DB 287 ACAGTGAGGAGGAGGAGGAGCATTGACAGCCACTGCTTGTCAAAAAGACCACTTG 346
QY 101 -----ArgProSerAlaProGlnI 107
DB 347 GGCAGAGATGAAGAGAGCCCAAGGAA 376

RESULT 9
PCT-US94-09789-1
; Sequence 1, Application PC/TUS9409789
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: NOVEL COMPONENT OF AMYLOID IN
; TITLE OF INVENTION: ALZHEIMER'S DISEASE AND METHODS FOR USE OF SAME
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East - Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09789
FILING DATE: 29-AUG-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: PD-3520
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1560 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: linear
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: cDNA for NACP
FEATURE:
NAME/KEY: misc RNA
LOCATION: 1..1560
PCT-US94-09789-1

Alignment Scores:
Pred. No.: 8.64e-30 Length: 1560
Score: 307.50 Matches: 68
Percent Similarity: 70.91% Conservative: 10
Best Local Similarity: 61.82% Mismatches: 29
Query Match: 50.41% Indels: 3
DB: 5 Gaps: 1

US-09-017-715A-2 (1-127) x PCT-US94-09789-1 (1-1560)
QY 1 MetAspValPheIySgIyPheSerIleAlaIySgIyValIglIyAlaIglu 20
DB 53 ATGGATGATATTCAAGAGCACTTCAAGGCCAAGGAGGAGTGTGGCTGCTGAG 112
QY 21 LyeThrIySgIyValIthrGluAlaIgluIyThrIySgIyValIgluIyMetTyVal 40
DB 113 AAAACCAACAGGCTGTGGCAGAGCAGGAAAGCAAAAGGCTGTCTATGTA 172
QY 41 G1AlaIyThrIySgIyValIthrGluAlaIgluIyThrIySgIyValIgluIyThrIyS 60
DB 173 GGCTCCAAACCAAGGAGGAGGAGTGTGATGCTGTGGCAACAGTGGCTGAGAAACCAA 232
QY 61 GluGlnAlaAsnAlaValSerIyAlaValIySerSerValIAsnThrValAlaThrIyS 80
DB 233 GAGCAAGTGAACAATGTTGAGAGCAGTGTGACGGGTGACAGCAGTAGCCAGAG 292
QY 81 ThyValIgluIyAlaIgluAsnIleAlaValIthrSerIyValIyAlaIygluIyAspIleu 100
DB 293 ACAGTGAGGAGGAGGAGGAGCATTGACAGCCACTGCTTGTCAAAAAGACCACTTG 352
QY 101 -----ArgProSerAlaProGlnI 107
DB 353 GGCAGAGATGAAGAGAGCCCAAGGAA 382

RESULT 10
US-09-949-016-3086
; Sequence 3086, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307

```



Db 418 ATGACGCGTTCATGATGAGGCGCTGTCATGCGCAAGAGGCGCTTGTGGCAGCGCGAG 477  
Qy 21 LysThrIysGlnGlyValThrGluAlaIaGluLysThrIysGlnGlyValMetTyVal 40  
Db 478 AAAACCAAGCGGGGGTCCACCGAGCGCGGAGGAGCAAGAGGAGGGGGTCTTACGTC 537  
Qy 41 G1ValAlaIysThrIysGlnGlyValIaGlnSerValThrSerValIaGlnLysThrIys 60  
Db 538 GGAAGCAAGACCCGAGAGGCTGTGTACAGCTGTGGCTTACGTGAGGAGGAGGAGGAG 597  
Qy 61 G1GluAlaIaSerValIaSerValIaValIaSerValIaSerValIaSerValIaThrIys 80  
Db 598 GAACAGGCGCTCACTCTGAGAGAGAGCTGTCTTCTGCGG----- 636  
Qy 81 ThrValGluGluAlaGluSerValIaValThrSerValIaValThrGluLys 99  
Db 637 -----GCAGGAGAACATCGACAGCAGCAGAGACTGTGAGAGAGAGAGAA 681

RESULT 13  
US-09-513-999C-2232  
; Sequence 2232, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.Y.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513.999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2232  
; LENGTH: 313  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 136..312  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 117  
; OTHER INFORMATION: m=a or c  
US-09-513-999C-2232

Alignment Scores:  
Pred. No.: 1.56e-27 Length: 313  
Score: 282.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.31% Mismatches: 0  
Query Match: 46.23% Indels: 0  
DB: 4 Gaps: 0

US-09-017-715A-2 (1-127) x US-09-513-999C-2232 (1-313)  
Qy 1 MetAspValPheIysLysGlyPheSerIleAlaIysLysGlyValIaGluValIaGlu 20  
Db 136 ATGATGCTTCAAGAGGGGCTTCTCATGCGCAAGAGGCGGTGTGGGCGGTGGA 195  
Qy 21 LysThrIysGlnGlyValThrGluAlaIaGluLysThrIysGlnGlyValMetTyVal 40  
Db 196 AAGGCCAAGCGGGGGTACCGAGAGCAGCTGAGAGAGCAAGAGAGGGGGCTATGTG 255  
Qy 41 G1ValAlaIysThrIysGlnGlyValIaGlnSerValThrSerValIaIaGluLysThr 59  
Db 256 GGAACCAAGACCCGAGAGGATGTGTACAGCGCTGAGCTGAGCGGAGAGAGACC 312

RESULT 14  
US-09-949-016-13657  
; Sequence 13657, Application US/09949016

; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C0001307  
; CURRENT APPLICATION NUMBER: US/09/949.016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13657  
; LENGTH: 8607  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(8607)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13657

Alignment Scores:  
Pred. No.: 4.69e-15 Length: 8607  
Score: 201.50 Matches: 97  
Percent Similarity: 20.66% Conservative: 3  
Best Local Similarity: 20.04% Mismatches: 3  
Query Match: 33.03% Indels: 383  
DB: 4 Gaps: 2

US-09-017-715A-2 (1-127) x US-09-949-016-13657 (1-8607)  
Qy 1 MetAspValPheIysLysGlyPheSerIleAlaIysLysGlyValIaGluValIaGlu 20  
Db 2048 ATGATGCTTCAAGAGGGGCTTCTCATGCGCAAGAGGCGGTGTGGGCGGTGGA 2107  
Qy 21 LysThrIysGlnGlyValThrGluAlaIaGluLysThrIysGlnGlyValMetTyVal 40  
Db 2108 AAGGCCAAGCGGGGGTACCGAGAGCAGCTGAGAGAGCAAGAGAGGGGGCTATGTG 2167  
Qy 40 ----- 40  
Db 2168 GGTAGTGGGCGATGCGAGGGGTGGACAGTGTGTGCGCAAGGGTGTAGCCAGTTAC 2227  
Qy 40 ----- 40  
Db 2228 CTTCGCAAGCTTACTTCCCGAGCCCGAGAGGAGGATTTGGAGGGGGCGAGCCCTGG 2287  
Qy 40 ----- 40  
Db 2288 CTATCAAGTGGGGTCTCAGAACCTTGAGCACCCACAAATGCCCTGTGCACTATGTG 2347  
Qy 40 ----- 40  
Db 2348 TGTTCCTTTGGCTTCCTCGGGGCTCTGTGGTGTGAGAGCGGAGAGGGCTGCTA 2407  
Qy 40 ----- 40  
Db 2408 CTGTCTGTGACAGCACACACATTCACAGATACAGCTTCCTGAGCTTGAGCC 2467  
Qy 40 ----- 40  
Db 2468 CTTGAAGCATGAGCAGCGCTGTCTCAGTGTGGCCCGACCTCTCCACAGGGAGGGCT 2527  
Qy 40 ----- 40  
Db 2528 ACAGCAGTCAAGGATCCCTCTCCCGAGAGAGAGGGGAGGCTGGGGATGAAC 2587  
Qy 40 ----- 40

```

Db      2588 CTAGGCTAGTGTCCCTCCCGGATCTCTCTGACACTCTCCAGAGAGAGAGGGA 2647
Qy      40 ----- 40
Db      2648 GGTCAAGCAATGACTCAGCTCTGGCCCATCTGTCCTGCTGCTGCTGAGCCGGC 2707
Qy      40 ----- 40
Db      2708 CACACCCGGGAGAGGGCTGGACCTGGGGCTAGACAGTCTCCCTACCTCAGGCTGCTCT 2767
Qy      40 ----- 40
Db      2768 CTCTGTCCCAACATCTGTCTCTGCTCCCTTCCATCCATCCATCTTTCAGACACAG 2827
Qy      40 ----- 40
Db      2828 CAGGAAGAGCCCTCTGAAGGGGCGCGCCGCCAGACACATCCTTACCCGCCACCG 2887
Qy      40 ----- 40
Db      2888 ACCCAACAGTTTGTCCAGCTGTTCTGTTGTTTGTCTGACCGCCCAACACCTCGAG 2947
Qy      41 ----- 49
Db      2948 GGAGGTCTGGGCTGACAGCTCATTTCTCCCAAGAGCCAGACCAAGAGATGTTGT 3007
Qy      49 |GlnSerVal|ThrSer----- 54
Db      3008 ACAGAGCGTGAACCTC-AGGTGAGAAAGCCCGAGGCGCAGGGGACACATGGGGATGAGACC 3066
Qy      54 ----- 54
Db      3067 CCTGGGGCTCTGCACTCTAGTGTGGGGCTCAAACTAGAGTCTGCTTACCCCAAC 3126
Qy      54 ----- 54
Db      3127 TGGGGTCCAGAGCCCTACAGACCCCTGACAGACCATGAGGCTAACTAGGGTGGGGCTTC 3186
Qy      54 ----- 54
Db      3187 CTTACCCCAACGATCAGAGGTGCTCTGAGTCAGAGGAGCAGGGGAGGGTCCACAC 3246
Qy      54 ----- 54
Db      3247 AAGGCCAGGGCTGTGAGCTCTGGAGAGGGGCTGCAGCCTGACTCAGACAGGCTGCCT 3306
Qy      55 ----- 60
Db      3307 TGGGGCTGGGGCTGGGGGTGAGGCGCAGCAGTGTCTCTCCCATAGTGGCCGAGAAACCAA 3366
Qy      60 |GlnGlnAlaAsnAlaValSerIleValValSerValAsnThrValAlaThrIle 80
Db      3367 GAGGCGGCGCAACGCGGTGAGGAGCTGTGGTGAACAGCTCAACATGTGGCCCA 3426
Qy      80 |GlnValGlnGlnAlaGlnAlaAsnIleAlaValThrSerGlnValValIleArgIleGlnAsp 100
Db      3427 GACCGTGGAGAGGCGGAGAACATCGCGGTCACTCCGGGTGTGGCGAAGGT-GAGCC 3485
Qy      100 |ArgProSer 103
Db      3486 CCGGCCCTCA 3495

```

```

RESULT 15
US-09-949-016-12184
; Sequence 12184, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016

```

```

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12184
; LENGTH: 8608
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(8608)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-12184

```

```

Alignment Scores:
Pred. No.: 4,696-15 Length: 8608
Score: 201.50 Matches: 97
Percent Similarity: 20.66% Conservative: 3
Best Local Similarity: 20.04% Mismatches: 3
Query Match: 33.03% Indels: 383
Gaps: 2

```

```
US-09-017-715A-2 (1-127) x US-09-949-016-12184 (1-8608)
```

```

Qy      1 MetAspValPheIleValGlyPheSerIleAlaIleValGlyValAlaValGln 20
Db      2048 ATGATGCTTCAAAAGAGGGCTTCCATCCGCAAGAGGGGCTGGTGGTGGTGA 2107
Qy      21 |LysThrIleGlnGlnValIleThrGlnAlaAlaGlnIleValIleValMetIleVal 40
Db      2108 AAGACCAAGCAGGGGGGTGCGGAAACAGCTGGAAGACCAAGAGGGGCTCATGTATGTG 2167
Qy      40 ----- 40
Db      2168 GGTAAGTGGGCGATGGCAGAGGTGGGACAGTGTGTGGCCAAAGGTGAGTCCAGTTAC 2227
Qy      40 ----- 40
Db      2228 CTTCGCAAGACTTACTCTCCCAAGCCCAAGAGGACATTTTGGAGGGGCGAGCCCTGG 2287
Qy      40 ----- 40
Db      2288 CTATCAAGGTGGGTCTCCAGACCTTGAGACCAACCAATGCTGTGCACTATGTG 2347
Qy      40 ----- 40
Db      2348 TGTGTGTCTTGGCTCTCGGGGGCTGTGGGTGCAGAGACCGAGACAGGGCTGGCTA 2407
Qy      40 ----- 40
Db      2408 CCTGTCTGACAGCAGACACACATTCCAAGCATACAGCCTCCCTGAGCCTGGAGCC 2467
Qy      40 ----- 40
Db      2468 CCTGAAGCCATGAGAGAGCTGTGTCTCAGGTGGCCCCACCTCTCCACAGGGAGCGCT 2527
Qy      40 ----- 40
Db      2528 ACAGCAGGTCAAGATCCCTCTCCCTCCCAAGAGAGAGGGGAGGCTGGGGATGAANC 2587
Qy      40 ----- 40
Db      2588 CTAGGCTAGTGTTCCTCCCGCATCTCTCTGACACTCTCCAGAGAGAGAGGGA 2647
Qy      40 ----- 40
Db      2648 GGTCAAGCAATGACTCAGCTCTGGCCCATCTGTCCTGTTGCTGAGGCCCGGC 2707

```



```

QY 40 ----- 40
Db 2708 CACACCCGGGAGGGGCTGAGACCTGGGTCTAGCCAGTGTCTTACTCAGAGCCTGCTCT 2767
QY 40 ----- 40
Db 2768 CTCTTGTCCCAATCTGTCTGTCCCTTCCATCATCACTTCTTCAGACAG 2827
QY 40 ----- 40
Db 2828 CAGAGAGGCCCTGTGAGGGGCGCGCCCAAGACACCATCTTACCCCAACCG 2887
QY 40 ----- 40
Db 2888 ACCCCACAGTTTGTCAAGCTGTCTGTGTGTCTGTCTGACCGCCCAACACTGAG 2947
QY 41 ----- 49
Db 2948 GAGGTCTGGGCTGACAGCTCATTTCTCCCAAGAGCCAAAGAGAAATGTTGT 3007
QY 49 |GlnSerValThrSer----- 54
Db 3008 ACAAGCGTGACTTC-AGGTGAGAGGCCAGGGCCAGGGGACATGGGGATAGAAC 3066
QY 54 ----- 54
Db 3067 CTGGGGCTCTGCATCTTGTGTCTGGGCTCAAACTAGAGTCTGCTTACCCCAAC 3126
QY 54 ----- 54
Db 3127 TGGGGTCCCAAGCCCTAAGACCCCTGAGACATGAGGCTAAACTAGGGTGGGCTCTC 3186
QY 54 ----- 54
Db 3187 CTTACCCCAACAGCATCAGAGTGCCTGTGAGTCAAGGAGCAGGGAGGTCCAGC 3246
QY 54 ----- 54
Db 3247 AGGGCCAGGGCTGTGAGCTCTGGGAAAGGGCTGCAAGCTGACTCCAGCAGGCTGCT 3306
QY 55 ----- 60
Db 3307 TGGGGCTGGGGCTGGGGTGAGGCCAGCAGTGTCTCCATAGTGGCCGAGAGACCAA 3366
QY 60 |GlnGlnAlaAsnAlaValSerIySaIaValSerSerValAsnThrValAlaThrIy 80
Db 3367 GAGACAGGCCCAACCGCTGAGCGCTGTGTGAGCAGCGCTCAACACTGGGCCACCAA 3426
QY 80 |SThrValGlnGlnAlaGlnAsnIleAlaValThrSerGlyValValArgLysGlnAspLe 100
Db 3427 GACCGTGGAGGAGGCGAGAACATCGCGTCACTCCGGGCTGGTGGCGCAAGGT-GAGCC 3485
QY 100 |uArgProSer 103
Db 3486 CCGGCCCTCA 3495

```

Search completed: May 4, 2005, 09:32:13  
 Job time : 305.06 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus2n model

Run on: May 4, 2005, 09:07:52 ; Search time 928.208 Seconds  
(without alignments)  
809.955 Million cell updates/sec

Title: US-09-017-715A-2

Perfect score: 610  
Sequence: 1 MDVFKGFGSIAXKGVGAVE.....EGEASKEKEVAEBAQSGSD 127

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool\_h/US9017715/runat\_04052005\_100743\_25600/app\_query.fasta\_1.661  
-DB=N\_Geneseq -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=ppct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ppct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US09017715\_@CGN\_1\_1\_703 @runat\_04052005\_100743\_25600 -ACPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_16Dec04:\*  
1: geneseq1980s:\*  
2: geneseq1990s:\*  
3: geneseq2000s:\*  
4: geneseq2001as:\*  
5: geneseq2001bs:\*  
6: geneseq2002as:\*  
7: geneseq2002bs:\*  
8: geneseq2003as:\*  
9: geneseq2003bs:\*  
10: geneseq2003cs:\*  
11: geneseq2003ds:\*  
12: geneseq2004as:\*  
13: geneseq2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the change being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	610	100.0	550	2	AAV42669 Human bre
2	610	100.0	550	3	AAI39470 Human HBG
3	610	100.0	550	6	ABL63343 Breast ca
4	610	100.0	550	6	ABV73813 Human gam
5	610	100.0	550	10	AAD63568 Human amy

6	610	100.0	550	10	ADG47636	Adg47636 Human amy
7	603	98.9	550	6	ABV73915	Abv73915 Human gam
8	602	98.7	783	4	AAI93778	AAI93778 Human pol
9	602	98.7	796	3	AAFI21784	AAFI21784 Human bre
10	595	97.5	488	12	ADM66887	Adm66887 Human hom
11	595	97.5	720	2	AAK29997	Aak29997 Human per
12	595	97.5	720	6	ABT76519	Abt76519 cDNA enco
13	592	97.0	720	10	ADBE3864	Adbe3864 Human SNC
14	592	97.0	720	12	ADH54342	Adh54342 Human SNC
15	538	88.2	479	9	ACH15493	Ach15493 Human adu
16	497	81.5	727	2	AAK29998	Aak29998 Mouse per
17	497	81.5	727	12	ADM66886	Adm66886 Murine ad
18	491.5	80.6	787	2	AAT51183	Aat51183 Human bre
19	468.5	76.8	786	9	ADA57841	Ada57841 Human CDN
20	464.5	76.1	990	13	ADR98806	Adr98806 Lung spec
21	328	53.8	210	4	AAD14356	Aad14356 Human gam
22	316	51.8	1018	10	ADBS2898	Adbs2898 Primary x
23	316	51.8	1018	10	ABT42008	Abt42008 Toxicity
24	316	51.8	1018	12	ADP71875	Adp71875 Renal tox
25	307.5	50.4	423	4	AAD14354	Aad14354 Human alp
26	307.5	50.4	423	6	AAD44409	Aad44409 Human alp
27	307.5	50.4	1105	8	AAL60583	Aal60583 GST-human
28	307.5	50.4	1466	6	ABZ35251	Abz35251 Human gen
29	307.5	50.4	1543	12	ADO24508	Ado24508 Human PRO
30	307.5	50.4	1543	12	ADP44624	Adp44624 Human alp
31	307.5	50.4	1560	2	AAQ83201	Aaq83201 Precursor
32	304.5	49.9	755	12	ADM86918	Adm86918 Human pro
33	304	49.8	441	8	ABX47454	Abx47454 Bovine ES
34	304	49.8	453	8	ABX39864	Abx39864 Bovine ES
35	303.5	49.8	1096	10	ADB37445	Adb37445 Human can
36	303.5	49.8	1096	12	ADP44633	Adp44633 Human alp
37	303.5	49.8	1228	4	AA802053	Aa802053 DNA encod
38	282	46.2	313	3	AAC02234	Aac02234 Human sec
39	253	41.5	249	4	AAD14355	Aad14355 Human bet
40	243	39.8	424	8	ABX36813	Abx36813 Bovine ES
41	233	38.2	473	9	ACH14620	Ach14620 Human adu
42	212.5	34.8	677	2	AAK04876	Aak04876 Human gam
43	205	33.6	456	9	ACH39765	Ach39765 Human foe
44	202	33.1	393	8	ABX38204	Abx38204 Bovine ES
45	201.5	33.0	4606	6	ABT10161	Abt10161 Human bre

#### ALIGNMENTS

RESULT 1	
AAV42669	
ID	AAV42669 standard; cDNA; 550 BP.
XX	
AC	AAV42669;
XX	
DT	09-NOV-1998 (first entry)
XX	
DE	Human breast cancer specific gene 1 (BCSG1) cDNA.
XX	
KW	Breast cancer specific gene 1; BCSG1; human; metastasis; diagnosis;
KM	therapy; genetic marker; ds.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
XX	12..395
XX	/*tag= a
XX	
PN	WO9833915-A1.
XX	
PD	06-AUG-1998.
XX	
PF	03-FEB-1998; 98WO-US001804.
XX	
PR	03-FEB-1997; 97US-0037080P.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	

PI Ji H, Rosen CA;  
XX  
XX MPI, 1998-446811/38.  
DR P-PSDB; AAW63123.  
XX  
XX New isolated human breast cancer specific gene - used to develop products  
PT for the diagnosis, clinical management and treatment of breast cancer and  
PT metastases.  
XX  
XX Claim 4; Fig 1; 73pp; English.  
XX  
XX This cDNA clone corresponds to the transcript of the newly identified  
CC human breast cancer specific gene 1 (BCSG1), and includes an open reading  
CC frame for a 14.2 kDa protein (see AAW63123). It was isolated from a  
CC breast cancer cDNA library following an EST search for novel genes  
CC differentially expressed in breast cancer versus healthy breast tissue.  
CC The clone is deposited at ATCC 97175 and ATCC 97856. A gradient and stage  
CC -specific BCSG1 expression has been demonstrated from virtually no  
CC detectable expression in normal or benign breast to low level and partial  
CC expression in low grade in situ breast carcinoma and high expression in  
CC infiltrating malignant breast carcinomas. BCSG1 is useful as a breast  
CC cancer progression marker. Recombinant vectors and host cells useful for  
CC recombinant production of BCSG1 polypeptides (including epitope-bearing  
CC polypeptides) are provided. BCSG1 polynucleotides, polypeptides and  
CC antibodies can be used for the detection of breast cancer cells or breast  
CC cancer metastasis, and to develop methods for the clinical management and  
CC treatment of breast cancer  
XX  
XX Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 6,79e-60 Length: 550  
Score: 610.00 Matches: 127  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
US-09-017-715A-2 (1-127) x AAW42669 (1-550)  
QY 1 Metaspvalphelylsyglpbeserillealalyeglyvalalglalvalglu 20  
DB 12 ATCGATGTTTCAAGAGGGCTTCTCCATCGCCCAAGAGGGGTGGTGGGAA 71  
QY 21 LysThrlyslnglylvalthrglualaalaglulvthrlyslnglylvalmetl 40  
DB 72 AGACCAACAGAGGGGTGACGAGAGAGCTGAGAGACCAAGAGGGGTCAATGATG 131  
QY 41 GYAlAlAlYsThrlYsGlUaSnVAlGlnSerValThrSeValAlaglulYsThrlYs 60  
DB 132 GGAGCCAAACCAAGAGAGATGTTGACAGAGGTCACTGAGCGGAGAGACCAAG 191  
QY 61 GIUGlnAlaAsnAlaValSerlysaAlaValSerSeValAsnThrValAlaThrly 80  
DB 192 GAGCAGGCCAACGCCCTGAGCAAGGCTGTGTGAGCAGGCTCAACCTGTGGCCAAAG 251  
QY 81 ThrValGlUGlnAlaGlnAsnIleAlaValThrSerGlyValAlaArglysluaPleu 100  
DB 252 ACCGTGAGAGAGCGGAGAAATCGCGGTCACTCCGGGGTGTGGCAAGAGACTTG 311  
QY 101 ATGProSeValAProGlnGlnGlnGlnGlnAlaSerlysluYsGlUglUValAlaGln 120  
DB 312 AAGGCATCTGCCCCCAACAGAGAGGTGAGGATCCAAAGAGAAAGAGAGTGGCAGAG 371  
QY 121 GIUAlaGlnSerlyGlyasp 127  
DB 372 GAGGCCAGAGTGGGAGAGAC 392  
RESULT 2  
AAA39470  
ID AAA39470 standard; DNA; 550 BP.  
XX  
XX AC AAA39470;

XX  
DT 24-AUG-2000 (first entry)  
XX  
XX Human HBGBA67A DNA.  
DE  
XX Human; ADA2; cytostatic; gene therapy; treatment; cancer;  
KW amyloid-like protein; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 12..395  
FT /\*tag= a  
FT /product= "HBGBA67"  
XX  
PN US6054289-A.  
XX  
XX 25-APR-2000.  
PD  
XX 30-AUG-1996; 96US-00705771.  
PF  
XX 30-AUG-1995; 95US-0002993P.  
PR  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
PI Moore PA;  
XX  
DR MPI, 2000-338491/29.  
DR P-PSDB; AAY87779.  
XX  
XX New polynucleotide encoding human AD2 is useful for treating cancer and  
PT for isolating cDNAs and genes having similar biological activity.  
PS Disclosure; Col 27-28; 54pp; English.  
XX  
XX This invention describes a novel polynucleotide (I) encoding human ADA2.  
CC The products of the invention have cytostatic activity and can be used  
CC for gene therapy. (II) is useful for treating cancer; as primers and  
CC probes for isolating full length cDNA and genes having similar biological  
CC activity. This sequence encodes a polypeptide derived from the human  
CC HBGBA67X clone which is an amyloid-like protein found in breast tissue  
XX  
SQ Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 6,79e-60 Length: 550  
Score: 610.00 Matches: 127  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
US-09-017-715A-2 (1-127) x AAA39470 (1-550)  
QY 1 Metaspvalphelylsyglpbeserillealalyeglyvalalglalvalglu 20  
DB 12 ATCGATGTTTCAAGAGGGCTTCTCCATCGCCCAAGAGGGGTGGTGGGAA 71  
QY 21 LysThrlyslnglylvalthrglualaalaglulvthrlyslnglylvalmetl 40  
DB 72 AGACCAACAGAGGGGTGACGAGAGAGCTGAGAGACCAAGAGGGGTCAATGATG 131  
QY 41 GYAlAlAlYsThrlYsGlUaSnVAlGlnSerValThrSeValAlaglulYsThrlYs 60  
DB 132 GGAGCCAAACCAAGAGAGATGTTGACAGAGGTCACTGAGCGGAGAGACCAAG 191  
QY 61 GIUGlnAlaAsnAlaValSerlysaAlaValSerSeValAsnThrValAlaThrly 80  
DB 192 GAGCAGGCCAACGCCCTGAGCAAGGCTGTGTGAGCAGGCTCAACCTGTGGCCAAAG 251  
QY 81 ThrValGlUGlnAlaGlnAsnIleAlaValThrSerGlyValAlaArglysluaPleu 100  
DB 252 ACCGTGAGAGAGCGGAGAAATCGCGGTCACTCCGGGGTGTGGCAAGAGACTTG 311

QY 101 ArgProSerAlaProGlnGlnGlyGluAlaSerIysGluGlyGlnGluValAlaGlu 120  
DB 312 AGGCGATCTGCCCCCAACAGAGGGTGGAGCATCCAAAGAAAGAGAAAGTGGCAGAG 371  
QY 121 GluAlaGlnSerGlyGlyAap 127  
DB 372 GAGGCCCAAGTGGGGGAGAGC 392  
RESULT 3  
ABLe63343  
ID ABL63343 standard; DNA; 550 BP.  
XX  
AC ABL63343;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Breast cancer related gene sequence SEQ ID NO:1680.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW Stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US010838.  
XX  
PR 05-JUN-2000; 2000US-0209473P.  
PR 05-JUN-2000; 2000US-0209531P.  
PR 18-SEP-2000; 2000US-023133P.  
PR 18-SEP-2000; 2000US-023161P.  
PR 20-SEP-2000; 2000US-023400P.  
PR 20-SEP-2000; 2000US-023403P.  
PR 20-SEP-2000; 2000US-023405P.  
PR 22-SEP-2000; 2000US-023450P.  
PR 22-SEP-2000; 2000US-023456P.  
PR 22-SEP-2000; 2000US-023492P.  
PR 25-SEP-2000; 2000US-023492P.  
PR 25-SEP-2000; 2000US-023507P.  
PR 25-SEP-2000; 2000US-023508P.  
PR 25-SEP-2000; 2000US-023513P.  
PR 25-SEP-2000; 2000US-023528P.  
PR 26-SEP-2000; 2000US-023563P.  
PR 26-SEP-2000; 2000US-023563P.  
PR 27-SEP-2000; 2000US-023571P.  
PR 27-SEP-2000; 2000US-023572P.  
PR 27-SEP-2000; 2000US-023584P.  
PR 27-SEP-2000; 2000US-023586P.  
PR 28-SEP-2000; 2000US-023602P.  
PR 28-SEP-2000; 2000US-023603P.  
PR 28-SEP-2000; 2000US-023603P.  
PR 28-SEP-2000; 2000US-023603P.  
PR 28-SEP-2000; 2000US-023610P.  
PR 28-SEP-2000; 2000US-023611P.  
PR 29-SEP-2000; 2000US-023684P.  
PR 29-SEP-2000; 2000US-023689P.  
PR 02-OCT-2000; 2000US-023717P.  
PR 02-OCT-2000; 2000US-023717P.  
PR 02-OCT-2000; 2000US-023729P.  
PR 02-OCT-2000; 2000US-023729P.  
PR 02-OCT-2000; 2000US-023731P.  
PR 02-OCT-2000; 2000US-023735P.  
PR 03-OCT-2000; 2000US-023758P.  
PR 03-OCT-2000; 2000US-023760P.  
PR 03-OCT-2000; 2000US-023760P.  
PR 03-OCT-2000; 2000US-023760P.  
PR 01-NOV-2000; 2000US-024486P.

PR 01-NOV-2000; 2000US-0245084P.  
XX  
XX (AVAL-) AVALON PHARM.  
XX  
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Sopet DR, Weaver Z;  
XX  
XX WPI; 2002-189264/24.  
XX  
PT Screening for anti-neoplastic agent involves exposing cells to a chemical  
PT agent to be tested for anti-neoplastic activity, and determining a change  
PT in expression of a gene of a signature gene set.  
XX  
PS Claim 1; SEQ ID NO 1680; 44pp; English.  
XX  
XX The present invention describes a method (M1) for screening for an anti-  
CC neoplastic agent. The method involves exposing cells to a chemical agent  
CC to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC activity and can be used in gene therapy. M1 can be used for screening an  
CC anti-neoplastic agent, and can be used for producing a product which is  
CC the data collected with respect to the anti-neoplastic agent as a result  
CC of M1, and the data is sufficient to convey the chemical structure and/or  
CC properties of the agent. M1 can be used in the treatment of cancer such  
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms'  
CC tumour.  
XX  
SQ Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;  
XX  
XX  
Alignment Scores:  
Pred. No.: 6 79e-60 Length: 550  
Score: 610.00 Matches: 127  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-09-017-715A-2 (1-127) x ABL63343 (1-550)  
QY 1 MetAspValPheIysIysGlyPheSerIleAlaIleIysGlyValValAlaIleValGlu 20  
DB 12 ATCGATGTTTCAAGAAAGGCTTCTCCATGCCAAGAGGCGTGGCGGTGAA 71  
QY 21 LysThrIysGlnGlyValThrGluAlaAlaGluIysThrIysGluValIleThrVal 40  
DB 72 AAGACCAAGCAGGGGGGAGCAGGAGCGTGAAGAGCCAAAGAGGGGGTCACTAATG 131  
QY 41 GluAlaIysThrIysGluAenValIaIleIysValIleSerValIleThrValIaIleThrIys 60  
DB 132 GAGGCCAAGACCAAGAGATGTTGTCAGAGGCTGAACCTCAAGTGGCCGGAAGACCAAG 191  
QY 61 GluGlnAlaAenAlaValSerIysAlaValIleSerSerValAenThrValIaIleThrIys 80  
DB 192 GAGCAGGCCAAGCCCGTAGACCAAGGCTGTGAGCAGCTCAACACTGTGGCCACCAAG 251  
QY 81 ThrValGluGluAlaGluAenIleAlaValIleSerGlyValValAlaIleGluAenPleu 100  
DB 252 ACCGTGAGAGAGCGGAGAACATCGCGTCACTCCGGGGTGTGCGCAAGAGGACTTG 311  
QY 101 ArgProSerAlaProGlnGlnGlyGluAlaSerIysGluGlyGlnGluValAlaGlu 120  
DB 312 AGGCGATCTGCCCCCAACAGAGGGTGGAGCATCCAAAGAAAGAGAAAGTGGCAGAG 371  
QY 121 GluAlaGlnSerGlyGlyAap 127  
DB 372 GAGGCCCAAGTGGGGGAGAGC 392

```

RESULT 4
ABV73813
ID ABV73813 standard; cDNA; 550 BP.
XX
XX ABV73813;
AC
XX 08-JAN-2003 (first entry)
DT
XX Human gamma-synuclein Glu110 variant gene.
DE
XX Gamma-synuclein; human; single nucleotide polymorphism; SNP;
KM schizophrenia; neuroleptic; gene; ss.
XX
XX Homo sapiens.
OS
FH
FH Key location/Qualifiers
FH CDS 12..395
FT /*tag= a
FT /product= "Gamma-synuclein"
FT /transl_except= (pos:213..215,aa:Glu)
FT replace(340,T)
FT variation /tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT
XX WO200275317-A2.
XX
XX 26-SEP-2002.
XX
XX 14-MAR-2002; 2002WO-EP002872.
XX
XX 15-MAR-2001; 2001US-0276306P.
XX
XX (NOVS ) NOVARTIS AG.
XX (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX (UYMA-) UNIV MARYLAND BALTIMORE.
XX
XX Roberts RC, Van Oostrum J, Voshol J, Tamminga CA;
XX WPI: 2002-750574/81.
XX DR P-PSDB; ABP54932.
XX
XX Screening for compounds for treating or interfering with the onset of
XX Schizophrenia Spectrum Disorders, by detecting interactions of candidate
XX compounds with the gamma-synuclein polypeptide.
XX
XX Disclosure; Fig 1: 32pp; English.
XX
XX The present sequence is that of cDNA encoding the Glu-110 isoform of
XX human gamma-synuclein. The invention relates to an isoform of gamma-
XX synuclein that is caused by an A/T single nucleotide polymorphism (SNP)
XX at position 329 of the gamma-synuclein coding sequence. This SNP causes a
XX glutamic acid to valine change at amino acid position 110 of gamma-
XX synuclein, and is associated with an increased susceptibility of
XX individuals to schizophrenia spectrum disorders (SSDs). This is the first
XX time that a genetic component of SSDs has been identified, and provides a
XX potential target for diagnosis and treatment of schizophrenia. Gamma-
XX synuclein polypeptides, especially those containing the E110V mutation,
XX are used in a claimed method of screening for compounds useful for the
XX treatment of SSDs, and gamma-synuclein expressing cells are used in a
XX claimed method of screening for agonist or antagonist compounds. An
XX oligonucleotide complementary to part of the gamma-synuclein coding
XX sequence is used for the discrimination of an SNP at position 329 of the
XX coding sequence. Gamma-synuclein polypeptides or polynucleotides are also
XX useful for the diagnosis of SSDs, or susceptibility to SSDs, e.g. by PCR
XX amplification of a polynucleotide encoding gamma-synuclein and analysis
XX of the occurrence of the SNP at position 329. A transgenic animal useful
XX for the study of SSDs is also claimed.
XX
XX Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;
XX
Alignment Scores: 6.79e-60 Length: 550
Pred. No.: 610.00 Matches: 127
Score:

```

```

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-017-715a-2 (1-127) x ABV73813 (1-550)
QY 1 MetAspValPheIysGlyPheSerIleAlaIysGlyValIglValAlaIglu 20
DB 12 ATGATGTTTTCAGAAAGGGCTTCCTCCATCCGCAAGAGGGCGTGTGGTCCGGTGGAA 71
QY 21 LysThrIysGlnGlyValIThrGluAlaIagIulYsThrIysGluGlyValMetIYrVal 40
DB 72 AAGACCAAGCAGGGGGGTACCGACAGCAGCTGGAAGACCAAGAGGGGGTCAATGATGTG 131
QY 41 GIYAlaIysThrIysGluAsnValValIInservAlThrServAlaIagIulYsThrIys 60
DB 132 GGAGCCAAAGACCAAGAGATGTTTTCACAGCGTACCTCGTGGCCGAGAGACCAAG 191
QY 61 GIUGlnAlaAsnAlaValSerIysAlaValIserSerValAsnThrValAlaThrIys 80
DB 192 GAGCAGGCCAAGCCGTGAGCAAGGCTGTGGTGAACAGGTCAACACTGTGGCCACCAAG 251
QY 81 ThrValIgluGluAlaGluAsnIleAlaValIThrSerGlyValIArgIysGluAsnIleu 100
DB 252 ACCGTGAGAGGCGGAGAACATCCGGTCACTCCGGGGTGTGTGGCAAGAGGACTTGG 311
QY 101 ArgProSerAlaProGlnIgluGluIasSerIysGluIysGluValAlaIglu 120
DB 312 AGGCCATCTGCCCCCAACAGAGGTGAGCATCCAAAGAAAGAAAGAACTGGCAGAG 371
QY 121 GluAlaGlnSerGlyIAsp 127
DB 372 GAGGCCCAAGTGGGGGAGAC 392
RESULT 5
AAD63568
ID AAD63568 standard; cDNA; 550 BP.
XX
XX AAD63568;
AC
XX 12-FEB-2004 (first entry)
DT
XX Human amyloid-like protein cDNA.
DE
XX Human; genetic disease; muscular dystrophy; cystic fibrosis; cytoskeletal;
XX scientific research; gene therapy; gene; amyloid-like protein; ss.
XX
XX Homo sapiens.
XX
XX Key location/Qualifiers
XX FH 12..395
XX FT CDS /*tag= a
XX FT /product= "Human amyloid-like protein"
XX
XX US6639052-B1.
XX
XX 28-OCT-2003.
XX
XX 14-OCT-1999; 99US-00417540.
XX
XX 30-AUG-1995; 95US-0002993P.
XX PR 30-AUG-1996; 96US-00705771.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Moore PA;
XX
XX WPI: 2003-842790/78.
XX DR P-PSDB; ABW02024.
XX
XX New isolated protein and nucleic acid molecules, useful for diagnostic
XX and therapeutic purposes, e.g. for treating genetic diseases such as

```

PT muscular dystrophy or cystic fibrosis.

XX Example 5; Fig 1; Opp; English.

XX  
CC The invention relates to isolated new isolated protein and nucleic acid  
CC molecules useful for diagnostic and therapeutic purposes. The invention  
CC is for treating genetic diseases such as muscular dystrophy or cystic  
CC fibrosis, and for in vitro purposes related to scientific research,  
CC synthesis of DNA and manufacture of DNA vectors. The invention is useful  
CC in gene therapy. The present sequence is human amyloid-like protein cDNA  
XX

SQ Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.79e-60 Length: 550  
Score: 610.00 Matches: 127  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-017-715A-2 (1-127) x AAD63568 (1-550)

QY 1 MetAspValPheLysGlyPheSerIleAlaLysGlyValAlaGlu 20

DB 12 ATGGATGTTTCAAGAGGGCTTCTCCATCCCAAGAGGGCGTGGCGGCGAA 71

QY 21 LysThrLysGlnGlyValThrGluAlaGluLysThrLysGlnGlyValMetTyrVal 40

DB 72 AAGCCAAAGAGGGGGTGAACGAGCAAGCTGAGAAAGCAAGAGGGGGTCACTATGTG 131

QY 41 GAlaLysThrLysGluAsnValAlaGlnSerValThrSerValAlaGluLysThrLys 60

DB 132 GGAAGCCAAAGCAAGAGAAAGTTGTACAGAGCGTGAACCTCAAGTGGCCGAAGAACCAAG 191

QY 61 GluGlnAlaAsnAlaValSerLysAlaValSerSerValAsnThrValAlaThrLys 80

DB 192 GAGCAGGCCAAAGCCCGTGAAGAGGCTGTGTGACAGCGCAACACTGTGGCCACCAAG 251

QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValAlaGlyLysGluAspLeu 100

DB 252 ACCGTGAGAGAGGGGAGAAACATCGCGTCACTCCGGGGTGTGGCCAAAGAGGACTTG 311

QY 101 ArgProSerAlaProGlnGlnGluGluAlaSerLysGluLysGluGluValAlaGlu 120

DB 312 AGGCCATCTCCGCCCAAGAGAGGTGAGGCATCAAGAGAAAGAGAAAGTGGCAGAG 371

QY 121 GluAlaGlnSerGlyLysAsp 127

DB 372 GAGGCCAGAGTGGGGAGAGAC 392

RESULT 6

ADG47636 ID ADG47636 standard; cDNA; 550 BP.

XX AC ADG47636;

XX 11-MAR-2004 (first entry)

DE Human amyloid like protein cDNA.

XX

XX ss: Gene; muscular dystrophy; cystic fibrosis; hypertension;

KM angina pectoris; myocardial infarction; ulcer; asthma; allergy;

KM psychosis; depression; migraine; vomiting; benign prostatic hypertrophy;

XX osteoporosis; human.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 12..395

FT /tag= a

FT /product= "Amyloid like protein"

PN US2003208043-A1.

XX 06-NOV-2003.

XX 04-JUN-2003; 2003US-00453478.

XX 30-AUG-1995; 95US-0002993P.

XX 30-AUG-1996; 96US-00705771.

XX 14-OCT-1999; 99US-00417540.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Moore PA, Gentz RL, Ji H, Ni J, Hu J;

XX WPI; 2003-864796/80.

XX P-PSDB; ADG47647.

XX New human polypeptides and polynucleotides, useful for diagnosing or

XX treating genetic diseases such as muscular dystrophy or cystic fibrosis,

XX hypertension, asthma, depression or osteoporosis.

XX Claim 18; SEQ ID NO 1; 56pp; English.

XX The invention relates to an isolated human polypeptide. The polypeptides,

XX polynucleotides, agonists or antagonist are useful for diagnosing or

XX treating genetic diseases such as muscular dystrophy or cystic fibrosis,

XX hypertension, angina pectoris, myocardial infarction, ulcers, asthma,

XX allergies, psychoses, depression, migraine, vomiting, benign prostatic

XX hypertrophy or osteoporosis. The polypeptides and polynucleotides are

XX useful for in vitro purposes related to scientific research, synthesis of

XX DNA and manufacture of DNA vector. The present sequence represents cDNA

XX encoding human amyloid like protein.

XX SQ Sequence 550 BP; 132 A; 145 C; 192 G; 81 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.79e-60 Length: 550  
Score: 610.00 Matches: 127  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-017-715A-2 (1-127) x ADG47636 (1-550)

QY 1 MetAspValPheLysGlyPheSerIleAlaLysGlyValAlaGlu 20

DB 12 ATGGATGTTTCAAGAGGGCTTCTCCATCCCAAGAGGGCGTGGCGGAA 71

QY 21 LysThrLysGlnGlyValThrGluAlaGluLysThrLysGlnGlyValMetTyrVal 40

DB 72 AAGCCAAAGAGGGGGTGAACGAGCAAGCTGAGAAAGCAAGAGGGGGTCACTATGTG 131

QY 41 GAlaLysThrLysGluAsnValAlaGlnSerValThrSerValAlaGluLysThrLys 60

DB 132 GGAAGCCAAAGCAAGAGAAAGTTGTACAGAGCGTGAACCTCAAGTGGCCGAAGAACCAAG 191

QY 61 GluGlnAlaAsnAlaValSerLysAlaValSerSerValAsnThrValAlaThrLys 80

DB 192 GAGCAGGCCAAAGCCCGTGAAGAGGCTGTGTGACAGCGTCAACACTGTGGCCACCAAG 251

QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValAlaGlyLysGluAspLeu 100

DB 252 ACCGTGAGAGAGGGGAGAAACATCGCGTCACTCCGGGGTGTGGCCAAAGAGGACTTG 311

QY 101 ArgProSerAlaProGlnGlnGluGluAlaSerLysGluLysGluGluValAlaGlu 120

DB 312 AGGCCATCTCCGCCCAAGAGAGGTGAGGCATCAAGAGAAAGAGAAAGTGGCAGAG 371

QY 121 GluAlaGlnSerGlyLysAsp 127

DB 372 GAGGCCAGAGTGGGGAGAGAC 392

RESULT 7  
 ABV73915  
 ID ABV73915 standard; cDNA; 550 BP.  
 XX  
 AC ABV73915;  
 XX  
 DT 08-JAN-2003 (first entry)  
 XX  
 DE Human gamma-synuclein Val110 variant gene.  
 KW Gamma-synuclein; human; single nucleotide polymorphism; SNP;  
 XX schizophrenia; neuroleptic; mutant; gene; ss.  
 OS Homo sapiens.  
 PH Key location/Qualifiers  
 FT CDS 12..395  
 FT /tag= a  
 FT /product= "Gamma-synuclein"  
 FT /transl\_except= (pos:213..215,aa:Glu)  
 FT replacement(340,A)  
 FT /tag= b  
 FT /standard\_name= "Single nucleotide polymorphism"  
 PN WO200275317-A2.  
 XX  
 XX 26-SEP-2002.  
 XX  
 XX 14-MAR-2002; 2002WO-EP002872.  
 XX  
 XX 15-MAR-2001; 2001US-0276306P.  
 XX  
 PA (NOVS ) NOVARTIS AG.  
 PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
 PA (UYMA-) UNIV MARYLAND BALTIMORE.  
 PI Roberts RC, Van Oostrum J, Voshol J, Tamminga CA;  
 XX  
 XX WPI: 2002-750574/81.  
 DR P-PSDB; ABB54933.  
 XX  
 PT Screening for compounds for treating or interfering with the onset of  
 PT Schizophrenia Spectrum Disorders, by detecting interactions of candidate  
 PT compounds with the gamma-synuclein polypeptide.  
 XX  
 PS Disclosure; Page; 32pp; English.  
 XX  
 CC The present sequence is that of cDNA encoding the Val-110 isoform of  
 CC human gamma-synuclein. The invention relates to an isoform of gamma-  
 CC synuclein that is caused by an A/T single nucleotide polymorphism (SNP)  
 CC at position 329 of the gamma-synuclein coding sequence. This SNP causes a  
 CC glutamic acid to valine change at amino acid position 110 of gamma-  
 CC synuclein, and is associated with an increased susceptibility of  
 CC individuals to schizophrenia spectrum disorders (SSDs). This is the first  
 CC time that a genetic component of SSDs has been identified, and provides a  
 CC potential target for diagnosis and treatment of schizophrenia. Gamma-  
 CC synuclein polypeptides, especially those containing the E110V mutation,  
 CC are used in a claimed method of screening for compounds useful for the  
 CC treatment of SSDs, and gamma-synuclein expressing cells are used in a  
 CC claimed method of screening for agonist or antagonist compounds. An  
 CC oligonucleotide complementary to part of the gamma-synuclein coding  
 CC sequence is used for the discrimination of a SNP at position 329 of the  
 CC coding sequence. Gamma-synuclein polypeptides or polynucleotides are also  
 CC useful for the diagnosis of SSDs, or susceptibility to SSDs, e.g. by PCR  
 CC amplification of a polynucleotide encoding gamma-synuclein and analysis  
 CC for the occurrence of the SNP at position 329. A transgenic animal useful  
 CC for the study of SSDs is also claimed. Note: The present sequence is not  
 CC shown in the specification but is derived from the gamma-synuclein  
 CC sequence given in Fig 1 (see ABV73913)  
 XX  
 SQ Sequence 550 BP; 131 A; 145 C; 192 G; 82 T; 0 U; 0 Other;  
 Alignment Scores:

Pred. No.:	4,226-59	Length:	550
Score:	603.00	Matches:	126
Percent Similarity:	99.21%	Conservative:	0
Best Local Similarity:	99.21%	Mismatches:	1
Query Match:	98.85%	Indels:	0
DB:	6	Gaps:	0

US-09-017-715A-2 (1-127) x ABV73915 (1-550)

QY	1 MetAapValPheLysGlyPheSerIleAlaLysGlyValValGlyAlaValGlu	20
DB	12 ATGATGTTTCAAGAGGGCTTCTCCATCGCCAAAGAGGGCTGGTGGTGGAA	71
QY	21 LysThrLysGlnGlyValThrGluAlaIleAlaGlyLysThrLysGluGlyValMetTyrVal	40
DB	72 AAGACCAACAGCGGGGTGACGGAACACAGTGAAGAACCAAGAGGGGTCTATGATGG	131
QY	41 GlyAlaLysThrLysGluAsnValValGlnSerValThrSerValAlaGluLysThrLys	60
DB	132 GGAGCCAAAGACCAAGAGATGTGTACAGACGCTGACCTCAGTGGCCGAGAACCAAG	191
QY	61 GluGlnAlaAsnAlaValSerLysAlaValLysSerValAsnThrValAlaThrLys	80
DB	192 GAGCAGGCCAAGCCGTGAGCAAGGCTGTGTGACACAGCTCAACACTGTGACCACAA	251
QY	81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValArgLysGluAspLeu	100
DB	252 ACCGTGAGAGAGCGGAGAACATCCGCTCACCTCCGGGTGTCCGCAAGAGACTTG	311
QY	101 ArgProSerAlaProGlnGlnGluGluAlaSerLysGluLysGluGluValAlaGlu	120
DB	312 AGGCCATCTGCCCCCAACAGAGGCTGTGCATCCAAAGAAAGAGAACTGGCAGAG	371
QY	121 GluAlaGlnSerGlyGlyAsp	127
DB	372 GAGGCCAAGATGGTGGAGAC	392

RESULT 8  
 AA193778  
 ID AA193778 standard; cDNA; 783 BP.  
 XX  
 AC AA193778;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 13838.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164835-A2.  
 XX  
 XX 07-SEP-2001.  
 XX  
 PF 26-FEB-2001; 2001WO-US004927.  
 XX  
 XX 28-FEB-2000; 2000US-00515126.  
 PR 18-MAY-2000; 2000US-00577409.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 XX WPI: 2001-514838/56.  
 DR P-PSDB; AAO13847.  
 XX  
 PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 PT and treating e.g. leukemia, inflammation and immune disorders.  
 XX



PS Claim 1; SEQ ID NO 13838; 1399bp + Sequence listing; English.  
XX  
CC The invention relates to human polynucleotides (AA179941-AA193841) and  
CC the encoded proteins (AA000010-AA01910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activity/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences  
XX  
SQ Sequence 783 BP; 187 A; 232 C; 237 G; 127 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 8,64e-59 Length: 783  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 4 Gaps: 0  
  
US-09-017-715A-2 (1-127) x AA193778 (1-783)  
QY 1 MetAspValPheIysGlyPheSerIleAlaIleYsgIyValaIglYalaIglu 20  
Db 114 ATGGATGCTTCAAGAGGGCTTCTCATGCCAAGAGGGCGGTGGGTGGAA 173  
QY 21 LysThrIysGlnGlyValThrGluAlaIglYlyThrIysGlnGlyValMetTYVal 40  
Db 174 AAGACCAAGAGGGGGTGGAGCAAGCGTGAAGACCAAGAGGGGGTCACTATGTG 233  
QY 41 GlYAlaIysThrIysGluAenValaIglNserValThrservAlaIgluYlThs 60  
Db 234 GGAGCCAGACCAAGAGGAATGTTGTACAGGCTGACCTCAGTGGCCGAGAACCAAG 293  
QY 61 GluGlnAlaAsnAlaValSerIysAlaIglValSerSerValAsnThrValaThrLys 80  
Db 294 GACGAGGCCAAGCGCGTGGAGCGGTGTGTGAGCGGCAACACTGTGGCCACCAAG 353  
QY 81 ThrValIgluGluAlaGluAenIleAlaValThrservYalValaIglYlThs 100  
Db 354 ACCGTGGAGAGCGGAGAACATCGCGTCACTCCGCGGTGTGGCCAGAGGACTTG 413  
QY 101 ArgProSerAlaProGlnGlnGluGluAlaSerIysGluYsgIgluValaIglu 120  
Db 414 AGGCCATCTGCCCCCAAGAGGGTGGAGGATCCAAAGAGAAAGTGGCAGAG 473  
QY 121 GluAlaGlnSerIglYlYAsp 127  
Db 474 GAGGCCAGAGTGGGGAGAC 494  
  
RESULT 9  
AA171784  
ID AAF21784 standard; DNA; 796 BP;  
XX  
AC AAF21784;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
XX Human breast and ovarian cancer associated antigen gene, SEQ ID 171.  
XX  
XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
KM neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
KM antidiabetic; antiinflammatory; antitumor; vulnery; anticonvulsant;  
KM antibacterial; antifungal; antiparasitic; cardiant; immune disorder;  
KM Addison's disease; allergy; autoimmune haemolytic anaemia;  
KM autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
KM multiple sclerosis; rheumatoid arthritis; ulcerative colitis;

KM cardiovascular disorder; wound healing; neurological disease; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO20005173-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 08-MAR-2000; 2000MO-US005881.  
XX  
XX 12-MAR-1999; 99US-0124270P.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Ruben SM,  
XX  
XX WPI; 2000-611515/58.  
DR P-PSDB; AAB58881.  
XX  
PT New human breast and ovarian cancer associated gene sequences and the  
PT polypeptides encoded by these genes, useful in the prevention, treatment  
PT and diagnosis of cancer, immune disorders, cardiovascular disorders and  
XX neurological diseases.  
XX  
PS Claim 1; Page 608; 1299bp; English.  
XX  
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
CC associated with breast and ovarian cancer. Included in the invention are  
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
CC isolation and characterization of the DNA and protein sequences of the  
CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
CC or antagonist sequences exhibit cytostatic; immunosuppressive; neotropic;  
CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;  
CC antiinflammatory; antitumor; vulnery; anticonvulsant; antibacterial;  
CC antifungal; antiparasitic and cardiant activity. The polynucleotide and  
CC protein sequences are used in the diagnosis of cancer, particularly  
CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists  
CC and agonists may also be used in the diagnosis, prevention and treatment  
CC of immune disorders e.g. Addison's disease, allergies, autoimmune  
CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC cardiovascular disorders such as myocardial ischemias; wound healing;  
CC neurological diseases such as cerebral anoxia and epilepsy; and  
XX infectious diseases  
SQ  
XX Sequence 796 BP; 195 A; 233 C; 240 G; 126 T; 0 U; 2 Other;  
  
Alignment Scores:  
Pred. No.: 8,82e-59 Length: 796  
Score: 602.00 Matches: 125  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 98.43% Mismatches: 0  
Query Match: 98.69% Indels: 0  
DB: 3 Gaps: 0  
  
US-09-017-715A-2 (1-127) x AAF21784 (1-796)  
QY 1 MetAspValPheIysGlyPheSerIleAlaIleYsgIyValaIglYalaIglu 20  
Db 109 ATGGATGCTTCAAGAGGGCTTCTCATGCCAAGAGGGCGGTGGGTGGAA 168  
QY 21 LysThrIysGlnGlyValThrGluAlaIglYlyThrIysGlnGlyValMetTYVal 40  
Db 169 AAGACCAAGAGGGGGTGGAGCAAGCGTGAAGACCAAGAGGGGGTCACTATGTG 228  
QY 41 GlYAlaIysThrIysGluAenValaIglNserValThrservAlaIgluYlThs 60  
Db 229 GGAGCCAGACCAAGAGGAATGTTGTACAGGCTGACCTCAGTGGCCGAGAACCAAG 288  
QY 61 GluGlnAlaAsnAlaValSerIysAlaIglValSerSerValAsnThrValaThrLys 80  
Db 289 GACGAGGCCAAGCGCGTGGAGCGGTGTGTGAGCGGCAACACTGTGGCCACCAAG 348

QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValAlaArgIleGluAspLeu 100  
Db 349 ACCGTGAGAGAGCCGAGAAATCGCGTCCGCGGTGTGTGGCAAGAGACTTG 408  
QY 101 ArgProSerAlaProGlnGlnGluGluAlaSerIleGluGluValAlaGlu 120  
Db 409 AGGCATCTGCCCCCAGACAGAGGTGAGCATCCAAAGAAAGAAAGTGGCAGG 468  
QY 121 GluAlaGlnSerGlyIleAsp 127  
Db 469 GAGGCCAGAGTGGGAGAGAC 489  
RESULT 10  
ID ADM66887 standard; DNA; 488 BP.  
AC ADM66887;  
XX  
XX 03-JUN-2004 (first entry)  
DE Human homologue of murine adipocyte specific gamma synuclein DNA Seq 20.  
XX  
XX human; adipocyte specific; gene; ds; adipose tissue; anti-obesity;  
KW high mobility group I-C protein; HMG1-C; obesity; leptin; ob; diabetes;  
KW adipogenesis; hypertension; cardiovascular disease; anorectic;  
KW antidiabetic; hypotensive; gamma synuclein.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2004011618-A2.  
XX  
XX 05-FEB-2004.  
XX  
XX 29-JUL-2003; 2003MO-US023684.  
XX  
XX 29-JUL-2002; 2002US-0398785P.  
XX PR 12-JUN-2003; 2003US-0478206P.  
XX  
XX (HMG1-) HMG1 INC.  
PA  
XX  
XX Chada K, Chouinard R, Ashar H, Sayed AMD;  
PI WPI: 2004-143846/14.  
XX  
XX P-PSDB; ADM67167.  
DR  
XX  
XX Identifying adipocyte specific genes, useful for treating obesity or  
PT diabetes, and for identifying drug targets, by differential gene  
PT expression analysis between adipose tissue or stromal vascular tissue of  
PT mice of different genotypes.  
XX  
XX Claim 11; SEQ ID NO 20; 91bp; English.  
XX  
XX This invention relates to a novel method for identifying genes that are  
XX over-expressed in adipose tissue and as such it provides targets for anti-  
XX -obesity pharmaceutical compositions. Specifically, it refers to a high  
XX mobility group I-C protein (HMG1-C) that is associated with obesity and  
XX is epistatic to leptin, furthermore, it refers to the ob gene where an  
XX autosomal recessive trait is linked to obesity and diabetes. The present  
XX invention describes performing differential gene expression analysis  
XX between the white adipose tissue (WAT) or stromal vascular tissue (SVT)  
XX of any two different mice selected from a group consisting of wild-type,  
XX HMG1-C -/-, ob/ob, or HMG1-C -/- ob/ob genotype mice. Accordingly, using  
XX this method novel nucleotides and the encoded proteins thereof were  
XX identified that are adipocyte specific, and as such can be used for  
XX preventing adipogenesis, diagnosing and treating diabetes, obesity,  
XX hypertension and cardiovascular disease, as well as screening for  
XX compounds that can prevent adipogenesis and treat diabetes or  
XX obesity. These compositions exhibit anorectic, antidiabetic and  
XX hypotensive activities. This polynucleotide sequence is a human homologue  
XX of a murine adipocyte specific DNA sequence of the invention.  
XX  
XX Sequence 488 BP; 127 A; 119 C; 176 G; 66 T; 0 U; 0 Other;  
XX  
XX

Alignment Scores:  
Pred. No.: 2,92e-58 Length: 488  
Score: 595.00 Matches: 124  
Percent Similarity: 99.21% Conservative: 2  
Best Local Similarity: 97.64% Mismatches: 1  
Query Match: 97.54% Indels: 0  
Gaps: 0  
US-09-017-715A-2 (1-127) x ADM66887 (1-488)  
QY 1 MetAspValPheIleValSerIleAlaIleValSerGlyValAlaGluValAlaGlu 20  
Db 12 ATGATGCTTTCACAAAGAGGCTTCCATCGCCAGAGAGGGGTGTGGATGCGGTGGA 71  
QY 21 LysThrIleGlnGlnValThrGluAlaAlaGluIleThrIleGlnGlnValMetIleVal 40  
Db 72 AAGACCAAGCAGGGGGGTGACGAGACAGCTGAGAAAGACCAAGAGGGGTCTATGTGTG 131  
QY 41 GluAlaIleThrIleGluAsnValAlaGlnSerValThrSerValAlaGluThrIle 60  
Db 132 GAGCCAAAGACCAAGAGAGATTTGTACAGACCTGACTCAGTGGCCGAGAAAGCCAAAG 191  
QY 61 GluAlaIleAsnAlaValSerIleValAlaValSerSerValAsnThrValAlaThrIle 80  
Db 192 GAGCAGGCGCAAGCCGTGAGCAGGCTGTGTGAGCAGAGCTCAACTGTGGCCACCAAG 251  
QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValAlaArgIleGluAspLeu 100  
Db 252 ACCGTGAGAGAGCCGAGAAATCGCGTCCGCGGTGTGTGGCAAGAGACTTG 311  
QY 101 ArgProSerAlaProGlnGlnGluGluAlaSerIleGluGluValAlaGlu 120  
Db 312 AGGCATCTGCCCCCAGACAGAGGTGAGCATCCAAAGAAAGAAAGTGGCAGAG 371  
QY 121 GluAlaGlnSerGlyIleAsp 127  
Db 372 GAGGCCAGAGTGGGAGAGAC 392  
RESULT 11  
ID AAX29997 standard; DNA; 720 BP.  
XX  
XX AAX29997;  
AC  
XX  
XX 06-JUL-1999 (first entry)  
DE Human persyn gene.  
XX  
XX Human persyn gene.  
KW Human; synuclein; persyn; diagnosis; neurodegenerative disorder; cancer;  
KW breast; skin; intermediate filament damage; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX EP908727-A1.  
XX  
XX 14-APR-1999.  
XX  
XX 21-SEP-1998; 98EP-00307628.  
XX  
XX 19-SEP-1997; 97GB-00019879.  
XX  
XX (NEUR-) NEUROPA LTD.  
XX (UYSA-) UNIV ST ANDREWS.  
PA  
XX  
XX WPI: 1999-217169/19.  
XX P-PSDB; AAY07271.  
DR  
XX  
XX New synuclein protein (persyn) and gene, useful in assays for screening,  
PT diagnosing or monitoring cancer, neurodegenerative disorders or skin  
PT disorders.  
XX  
XX Claim 29; Page 16-17; 39bp; English.  
XX  
XX



DB 169 GGAGCCAGACCAAGAGATGTTGTACAGAGGTGACCTCACTGGCCGAGAAAGCAAG 228  
QY 61 GUGGlnAlaAsnAlaValSerIySaIValSerSerValAsnThrValAlaThrIyS 80  
DB 229 GAGCAGGCCCAAGCGGTGAGCGAGGTGTGTGAGAGCGGTCAACACTGTGCCCACCAAG 288  
QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValArgIySgIuAspIeu 100  
DB 289 ACCGTGAGAGAGCGGAGAAACATCGCGTCACTCCGGGTGTGTGCGCAAGAGACTTG 348  
QY 101 ArgProSerAlaProGlnGlnGluGluAlaSerIySgIuIySgIuValAlaGlu 120  
DB 349 AGCCCATCTGCCCGCCCAACAGAGGGGTGTGTGATCCAAAGAGAAAGAGAGTGTGCAGAG 408  
QY 121 GIuAlaGlnSerGlyIyAsp 127  
DB 409 GAGGCCAGAGTGGGGAGAC 429  
RESULT 13  
ADE43864  
ID ADE43864 standard; cDNA, 720 BP.  
XX ADE43864;  
XX 29-JAN-2004 (first entry)  
XX Human SNGC cDNA, SEQ ID 469.  
DE Human SNGC cDNA, SEQ ID 469.  
XX Neurodegenerative disease; uPA; SNGC; IDE; KNSL1; LIPA; TNFRSF6;  
KW Alzheimer's disease; neuroprotective; neurotrophic; gene therapy;  
XX Chromosome 10; gene; ss.  
XX Homo sapiens.  
XX WO2003054143-A2.  
XX 03-JUL-2003.  
XX 25-OCT-2002; 2002WO-US034679.  
XX 25-OCT-2001; 2001US-0339525P.  
XX 08-NOV-2001; 2001US-0336929P.  
XX 08-NOV-2001; 2001US-0338010P.  
XX 09-NOV-2001; 2001US-0338363P.  
XX 04-DEC-2001; 2001US-0337052P.  
XX 28-MAR-2002; 2002US-0368919P.  
XX (NEUR-) NEUROGENETICS INC.  
XX (GENO-) GEN HOSPITAL CORP.  
XX Becker KD, Velicelch G, Elliott KJ, Wang X, Tanzi RE, Bertram L;  
XX Saunders AJ, Mullin KM, Sampson AJ, Blacker DL;  
XX WPI; 2003-559131/52.  
XX Determining a predisposition for or the occurrence of neurodegenerative  
XX disease, e.g. Alzheimer's disease by detecting in a target nucleic acid  
XX the presence or absence of an allelic variant of one or more polymorphic  
XX regions.  
XX Claim 84; Page 740; 848p; English.  
XX The present invention relates to a method (M1) for determining a  
XX predisposition for or the occurrence of neurodegenerative disease in a  
XX subject. The method comprises detecting in a target nucleic acid obtained  
XX from the subject the presence or absence of an allelic variant of one or  
XX more polymorphic regions of one or more genes selected from uPA  
XX (urokinase plasminogen activator), SNGC (gamma-synuclein), IDE (insulin-  
XX degrading enzyme), KNSL1 (Kinesin-like protein 1), LIPA (lysosomal acid  
XX lipase), and TNFRSF6 (Tumour Necrosis Factor Receptor-SF6), where the  
XX presence of at least one of the allelic variant of one or more  
XX polymorphic regions is indicative of a predisposition for or the  
XX occurrence of neurodegenerative disease. The genes are all located on

CC chromosome 10. M1 is useful for determining a predisposition for or the  
CC occurrence of, and for treating neurodegenerative disease, particularly  
CC Alzheimer's disease.  
XX  
XX SQ Sequence 720 BP; 172 A; 205 C; 212 G; 120 T; 0 U; 11 Other;  
Alignment Scores:  
Pred. No.: 1,056-57 Length: 720  
Score: 592.00 Matches: 124  
Percent Similarity: 97.64% Conservative: 0  
Best Local Similarity: 97.64% Mismatches: 3  
Query Match: 97.05% Indels: 0  
DB: 10 Gaps: 0  
US-09-017-715A-2 (1-127) x ADE43864 (1-720)  
QY 1 MetAspValPheIyAlSerGlyPheSerIleAlaIySgIyValIyAlIyValIyAlIy 20  
DB 49 ATGATGTTTTCAGAAAGGGCTTCTCCATCCGCAAGNAGGGCGTGTGTGCGGTGGA 108  
QY 21 LysThrIySgIyValIyThrGluAlaIyIySgIyThrIySgIyValIyMetIyVal 40  
DB 109 AAGACCAAGCAGGGGGTGAAGAGACAGCTGAAAGACCAAGAGGGGTGATGTATGTG 168  
QY 41 GIyAlaIyThrIySgIyAsnValIyGlnSerValIyThrSerValIyAlIyIySgIyThrIyS 60  
DB 169 GGAGCCAGACCAAGAGAGATTTGTACAGACGTGACCTGCTGTCGCGAGAAACCAAG 228  
QY 61 GUGGlnAlaAsnAlaValSerIySaIValSerSerValAsnThrValAlaThrIyS 80  
DB 229 GAGCAGGCCCAAGCGGTGAGCGAGGTGTGTGAGAGCGGTCAACACTGTGCCCACCAAG 288  
QY 81 ThrValGluGluAlaGluAsnIleAlaValThrSerGlyValValArgIySgIuAspIeu 100  
DB 289 ACCGTGAGAGAGCGGAGAAACATCGCGTCACTCCGGGTGTGTGCGCAAGAGACTTG 348  
QY 101 ArgProSerAlaProGlnGlnGluGluAlaSerIySgIuIySgIuValAlaGlu 120  
DB 349 AGCCCATCTGCCCGCCCAACAGAGGGGTGTGTGATCCAAAGAGAAAGAGAGTGTGCAGAG 408  
QY 121 GIuAlaGlnSerGlyIyAsp 127  
DB 409 GAGGCCAGAGTGGGGAGAC 429  
RESULT 14  
ADH54342  
ID ADH54342 standard; cDNA, 720 BP.  
XX ADH54342;  
XX 25-MAR-2004 (first entry)  
XX Human SNGC gene cDNA sequence SeqID469.  
XX human; neurodegenerative disease; urokinase plasminogen activator; uPA;  
XX gamma-synuclein; SNGC; insulin degrading enzyme; IDE;  
XX kinesin-like protein 1; KNSL1; lysosomal acid lipase; LIPA;  
XX tumour necrosis factor receptor SF6; TNFRSF6; Alzheimer's disease; ss.  
XX Homo sapiens.  
XX US2003224380-A1.  
XX 04-DEC-2003.  
XX 25-OCT-2002; 2002US-00282174.  
XX 25-OCT-2001; 2001US-0339525P.  
XX 25-OCT-2001; 2001US-0348065P.  
XX 02-NOV-2001; 2001US-0336983P.  
XX 08-NOV-2001; 2001US-0336929P.  
XX 08-NOV-2001; 2001US-0338010P.  
XX 09-NOV-2001; 2001US-0338363P.

PR 04-DEC-2001; 2001US-0337052P.  
PR 26-MAR-2002; 2002US-0368919P.  
XX (GEHO) GEN HOSPITAL CORP.  
XX Becker KD, Velicelab G, Elliott KJ, Wang X, Tanzi RE;  
PI Bertram L, Saudere AJ, Mullin KM, Sampson AJ;  
XX WPI; 2004-060538/06.  
XX  
XX Determining a predisposition for or the occurrence of neurodegenerative  
PT disease, particularly Alzheimer's disease, comprises determining the  
PT presence of a polymorphism in the upa, SNCG, IDE, KNSL1, LIPA or TNFRSF6  
PT gene.  
XX  
XX Claim 84; SEQ ID NO 469; 205bp; English.  
XX  
XX This invention relates to a novel method of determining a predisposition  
CC for or the occurrence of neurodegenerative disease comprising detecting  
CC in a target nucleic acid obtained from the subject the presence of an  
CC allelic variant of polymorphic regions of human genes selected from  
CC urokinase plasminogen activator (uPA), gamma-synuclein (SNCG), insulin  
CC degrading enzyme (IDE), kinesin-like protein 1 (KNSL1), lysosomal acid  
CC lipase (LIPA) and tumour necrosis factor receptor SF6 (TNFRSF6). The  
CC method is useful in determining the presence or predisposition to a  
CC neurodegenerative disease, particularly Alzheimer's disease. The present  
CC sequence is the cDNA sequence of the human SNCG gene which is related to  
CC the invention.  
XX  
XX Sequence 720 BP; 172 A; 205 C; 212 G; 120 T; 0 U; 11 Other;  
SQ  
XX  
XX Alignment Scores:  
Pred. No.: 1.05e-57 Length: 720  
Score: 592.00 Matches: 124  
Percent Similarity: 97.64% Conservative: 0  
Best Local Similarity: 97.64% Mismatches: 3  
Query Match: 97.05% Indels: 0  
Gaps: 0  
DB: 12  
US-09-017-715A-2 (1-127) x ADH54342 (1-720)  
QY 1 MetAspValPheLysGlyPheSerIleAlaLysGlyValAlaGlyAlaGlu 20  
Db 49 ATGATTTTTCAGAAAGGCTTCTCATCCCAAGAAAGGCGTGTGGCGGTGAA 108  
QY 21 LysThrLysGlnGlyValThrGluAlaAlaGluLysThrLysGlnGlyValMetTyrVal 40  
Db 109 AAGACCAAGCAGGGGGTGACGGAAGCAGCTGAGAAAGCAAGAGGGGGTATATGTG 168  
QY 41 GAlaAlaLysThrLysGluLysValAlaGlnSerValThrSerValAlaGluLysThrLys 60  
Db 169 GGAGCCAAAGCAGGAGAAATGTTGTACAGGCTGTGACTCAGTGGCCGGAACCAAG 228  
QY 61 GAlaGlnAlaSerLysValSerLysValValSerSerValAenThrValAlaThrLys 80  
Db 229 GAGCAGGCCCAACGNGAGAGCAGGCTGTGTGAGCGCTCAACACTGTGGCCACCAAG 288  
QY 81 ThrValGlnGluAlaGluAenIleAlaValThrSerGlyValValAlaGlyLysPleu 100  
Db 289 ACCGTGAGAGGGGAGGAACATCGCGTCACTCCGGGGGTGGCCCAAGAGACTTG 348  
QY 101 ArgProSerLysProGlnGlnGlnGlyGluAlaSerLysGlnGlnGluAlaGlu 120  
Db 349 AGGCCCATCTTCCCTCCCAAGAGAGGTGNGCATCCAAAGAAAGAGAAAGTGGCAGAG 408  
QY 121 GAlaAlaGlnSerGlyGlyAsp 127  
Db 409 GAGGCCCAAGTGGGGAGAC 429  
RESULT 15  
ACH15493  
XX ACH15493 standard; cDNA; 479 BP.  
XX

AC ACH15493;  
XX 13-OCT-2003 (first entry)  
XX Human adult brain cDNA #2705.  
XX  
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
XX genome mapping; biodiversity; genetic disorder.  
XX  
XX Homo sapiens.  
XX  
XX US2003073623-A1.  
XX  
XX 17-APR-2003.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX (DRMA/) DRMANAC R T.  
XX (LABA/) LABAT I.  
XX (STAC/) STACHE-CRAIN B.  
XX (DICK/) DICKSON M C.  
XX (JONE/) JONES L W.  
XX  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 2705; 44bp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations,  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?docid=20030073623  
XX  
SQ Sequence 479 BP; 120 A; 118 C; 178 G; 56 T; 0 U; 7 Other;  
SQ  
XX  
XX Alignment Scores:  
Pred. No.: 8.23e-52 Length: 479  
Score: 538.00 Matches: 112  
Percent Similarity: 99.14% Conservative: 3  
Best Local Similarity: 96.55% Mismatches: 1  
Query Match: 88.20% Indels: 0  
Gaps: 0  
DB: 9  
US-09-017-715A-2 (1-127) x ACH15493 (1-479)  
QY 12 LysLysGlnValAlaGluLysThrLysGlnGlyValThrGluAlaGlu 31  
Db 47 CAGAGGGCGCTGGGCTGGGAGAAAGACCAAGAGGGGGTGACGGAAGCACTGAG 106  
QY 32 LysThrLysGlnGlyValMetTyrValGlyAlaLysThrLysGlnGluLysValGlnSer 51  
Db 107 AAGACCAAGAGGGGGTCACTGTATGTGGAGCCCAAGCAAGAGAAATGTTGTACAGAC 166

```
QY 52 ValThrSerValAlaGluIysThrIysGluGlnAlaAsnAlaValSerIysAlaValVal 71
Db 167 GTGACCTCAGTGGCCGAGAGACCAAGGACCAAGCCCGTGGAGGAGGCTGTGGTG 226
QY 72 SerSerValAsnThrValAlaThrIysThrValGluGluAlaGluAsnIleAlaValThr 91
Db 227 AGCAGCGTCACACCTGTGGCCACCAAGACCCTGTGAGAGAGCGGAGAACATCGCGGTCAAC 286
QY 92 SerGlyValValAlaArgIysGluAspLeuArgProSerAlaProGlnGlnGluGluAla 111
Db 287 TCCGGGGGTGTGGCAGAGGAGACTTGAAGCCATCTGCCCCCAACAGAGAGGTGAGGCA 346
QY 112 SerIysGluIysGluGluValAlaGluGluAlaGlnSerGlyValYasp 127
Db 347 TCCAAAGAGAAAGAGAGAGTGGCAGAGAGGCCCAAGAGTGGGGGAGAGC 394
```

Search completed: May 4, 2005, 09:26:16  
Job time : 934.208 secs